NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

A MULTIVARIATE TIME SERIES ANALYSIS OF U.S. ARMY RECRUITING

by

Eric C. Burger

June 2000

Thesis Advisor: Samuel E. Buttrey
Second Reader: Lyn R. Whitaker

Approved for public release; distribution is unlimited

20000720 022

REPORT DOCUMENTATION PAGE		*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.		
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** June 2000	**3. REPORT TYPE AND DATES COVERED** Master's Thesis
TITLE AND SUBTITLE : A Multivariate Time Series Analysis Of U.S. Army Recruiting		**5. FUNDING NUMBERS**
6. AUTHOR(S) Eric C. Burger		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000		**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A		**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.		
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.		**12b. DISTRIBUTION CODE** A

ABSTRACT ***(maximum 200 words)***

The United States Army Recruiting Command requires tools to quantify the impact of factors in the recruiting environment, identify differences in the recruiting processes across its five regional subordinate units, and measure the effectiveness of its policies and resource expenditures. This thesis examines recruiting data for the "high-quality" male demographic from July 1992 to September 1997. It uses multivariate time series analysis to predict the number of enlistment contracts signed in a month as a function of fifteen exogenous and endogenous factors plus monthly indicators. A stepwise recursion using bootstrap simulation is developed to identify significant factors in the multivariate time series. The significant factors in the reduced models are compared to those contained in models developed in previous studies. The models are also used to create nine-month projections of recruiting production, which are compared to known production figures from test set data to determine forecast accuracy. The results of this research support the intuition that the influential factors differ by region. The stepwise model reduction recursion using bootstrap simulation offers potential for further refinement and application.

14. SUBJECT TERMS Time Series, Bootstrap, Recruiting, Operations Research			**15. NUMBER OF PAGES** 114
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

A MULTIVARIATE TIME SERIES ANALYSIS OF U.S. ARMY RECRUITING

Eric C. Burger
Captain, United States Army
B.S., United States Military Academy, 1990

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 2000

Author: Eric C. Burger
Eric C. Burger

Approved by: Samuel E. Buttrey
Samuel E. Buttrey, Thesis Advisor

Lyn R. Whitaker
Lyn R. Whitaker, Second Reader

Richard E. Rosenthal
Richard E. Rosenthal, Chairman
Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The United States Army Recruiting Command requires tools to quantify the impact of factors in the recruiting environment, to identify differences in the recruiting processes across its five regional subordinate units, and to measure the effectiveness of its policies and resource expenditures. This thesis examines recruiting data for the "high-quality" male demographic from July 1992 to September 1997. It uses multivariate time series analysis to predict the number of enlistment contracts signed in a month as a function of fifteen exogenous and endogenous factors plus monthly indicators. A stepwise recursion using bootstrap simulation is developed to identifying significant factors in the multivariate time series. The significant factors in the reduced models are compared to those contained in models developed in previous studies. The models are also used to create nine-month projections of recruiting production, which are compared to known production figures from test set data to determine forecast accuracy. The results of this research support the intuition that the influential factors differ by region. The stepwise model reduction recursion using bootstrap simulation offers potential for further refinement and application.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS

ACF	Autocorrelation Function
AFQT	Armed Forces Qualification Test
AIC	Akaike Information Criterion
AR	Autoregressive
ARMA	Autoregressive Moving Average
ASVAB	Armed Services Vocational Aptitude Battery
CLEMM	Command Level Mission Model
DCSPER	Deputy Chief of Staff for Personnel
DMDC	Defense Manpower Data Center
DOD	Department of Defense
FIPS	Federal Information Processing Standards
FY	Fiscal Year
GED	General Equivalency Diploma
MA	Moving Average
MEPS	Military Entrance Processing Station
ODCSPER	Office of the Deputy Chief of Staff for Personnel
TSC	(ASVB) Test Score Category
USAREC	U.S. Army Recruiting Command
YATS	Youth Attitude Tracking Survey

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

The United States Army is experiencing the greatest recruiting shortages since the inception of the All-Volunteer Force in 1973. The service faces unprecedented competition for young people as unemployment is at its lowest level in thirty years and college attendance rates are the highest in American history. The U.S. Army Recruiting Command (USAREC) is the organization charged with recruiting civilians for service in the Army. USAREC requires tools to quantify the impact of factors in the recruiting environment, identify differences in the recruiting processes across its five regional subordinate units, and measure the effectiveness of its policies and resource expenditures. This thesis examines recruiting data for from July 1992 to September 1997, which was a very dynamic period for the Army and Recruiting Command. The scope is limited to the high-quality male demographic. The Army defines a high-quality recruit as one who scored above the 50th percentile on the Armed Forces Qualification Test and who is a high school graduate or general equivalency diploma holder.

A considerable amount of research has been dedicated to the topic of Army recruiting. One of the goals of this thesis is to validate factors from previous models on more current data. Many observers have proposed new or changing influences on the recruiting environment. A further objective of this thesis is to explore these suppositions quantitatively by combining new factors with ones previously shown to be significant. Enumeration of the differences in the recruiting environment throughout the country is another objective. Finally, this thesis aims to develop an accurate tool for predicting recruiting production that can be used by Army leaders.

Multivariate time series analysis is used to predict the number of enlistment contracts signed in a month as a function of exogenous and endogenous factors plus monthly indicators. Fifteen factors are initially included for examination in this study as predictive variables. They are selected based on their appearance in previous models or in recent research. Autoregressive moving average (ARMA) models are developed to produce residuals with a suitable structure for bootstrapping. The bootstrap is used to overcome the difficulties in determining significant factors presented by the short

duration of the recruiting data time series. This technique allows resampling from within the existing data to provide robustness in the factor determination process. A stepwise recursion is developed to eliminate factors from the time series models that are not statistically significant. The factors remaining in the reduced models are compared to those found to be significant in past research. The developed models are also used to create nine-month projections of recruiting production. The results are then compared to known production figures from test set data to determine forecast accuracy levels.

The final models indicate that unemployment figures and high school graduate wage levels are significant factors for predicting recruiting production. These results are consistent with findings from previous studies. However, the impact of these two factors is not clearly interpretable across the five recruiting brigades. No consistent factors for measuring the competition between the Army and post-secondary schooling emerge from the model development process. The final models do successfully capture the seasonal nature of recruiting. There are considerable differences in the final model for each brigade, indicating that influential predictors of recruiting production differ regionally. The forecasts produced using the final models capture the general behavior of the recruiting production series in the test period. The stepwise recursion using bootstrap simulation for identifying significant factors in multivariate time series analysis proved to be a useful tool and offers potential for further refinement and application.

ACKNOWLEDGEMENTS

I would like to express my thanks to Professor Sam Buttrey for his problem-solving tenacity and patient teaching abilities, and to Professor Lyn Whitaker for her valued insight and input as a second reader. I am grateful for the support of Mr. Robert Tinney of the Defense Manpower Data Center and Dr. John Warner of Clemson University, who provided the data for this thesis. I would also like to recognize Mr. Dennis Mar of the Naval Postgraduate School Systems Management Department who generously offered his time and expertise to aggregate voluminous data into a useable form for analysis. Finally, I would like to express my appreciation to my wife, Christie, for her understanding, support, and encouragement.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

The United States Army is currently experiencing the greatest recruiting shortages since the inception of the All-Volunteer Force. The service faces many challenges in the recruiting realm. Competition for young people is unprecedented as unemployment has recently reached 30-year lows. The nation's youth have demonstrated a decreasing propensity to enlist as measured by the annual Youth Attitude Tracking Survey (YATS) and have enjoyed the highest college attendance rates in U.S. history (Parlier, 1999). As a result of these and other factors, the Army has failed to meet its recruiting requirements every year since 1997.

The current recruiting conditions are in stark contrast to those of the early 1990's, a period that represented unequalled success in terms of the quantity and quality of soldiers recruited by the Army. This achievement coincided with a decreased recruiting demand as the active Army force was pared down from its cold war level of close to 750,000 to its current strength of approximately 480,000. As the force was reduced, recruiting requirements decreased 33% (Asch, 1999). Towards the conclusion of the drawdown, the Army entered a "steady state," meaning that every soldier who left the service had to be replaced by a new recruit. Hence, accession requirements have actually increased slightly since 1995.

In response to recent shortcomings in meeting recruiting objectives, The Army Chief of Staff, General Shinseki, declared recruiting "the number one mission on his essential task list" (Dickey, 1999). The organization charged with the mission of recruiting civilians for service in the Army is the U.S. Army Recruiting Command (USAREC). To improve performance, USAREC has increased recruiter strength 15% since the beginning of 1997. It is also offering costly new enlistment incentives.

USAREC is dedicated to matching people to Army personnel requirements. Like many high-tech organizations, the Army seeks to fill its ranks with "high-quality

recruits." The Army defines a high-quality recruit as one who scored above the 50th percentile on the Armed Forces Qualification Test (AFQT) and who is a high school graduate or general equivalency diploma (GED) holder. Army policies over the past decade have required that 90 to 95% of all accessions have a high school diploma or GED. Since there are such demanding policy requirements for high quality recruits, this demographic category receives a majority of focus and recruiting effort.

B. STATEMENT OF PROBLEM

The troubled status of recruiting has gained public attention. The challenges in this arena are well documented and USAREC is applying more resources to achieve its objectives. Simple allocation of greater funds to USAREC alone is not the answer to the service's manpower shortcomings. Precise application of these monies is critical. As pointed out by RAND researcher Bruce Orvis, "decisions about increases [must be] preceded by identification of specific shortages that need to be remedied" (Orivs, 1996). In a period of increasing competition for eligible recruits, the Army's challenges will not recede in the foreseeable future. Therefore, USAREC must operate with the greatest possible efficiency in its application of limited resources.

Under these conditions USAREC requires tools to measure the effectiveness of its policies and resource expenditures and to apply an appropriate balance of effort across its five major subordinate units, which represent geographical regions of the United States. This thesis uses multivariate time series analysis to predict recruiting production (the number of enlistment contracts signed in a given period) as a function of exogenous and endogenous factors.

1. Research Questions

Models that address macro-level policies, recruiter distribution, and allocation of resources have utility for USAREC. The following research questions motivated the development of the time series models in this thesis:

a. What are the most significant economic, demographic, and policy predictors of recruiting success?

b. What are the differences between the five regional recruiting brigades regarding these various factors?

c. How effectively can recruiting production be predicted using a multivariate time series model?

2. Scope and Assumptions

The models developed in this thesis predict production at the regional level. The scope of this study is limited to the high-quality male demographic, which single largest category of recruits accessed. Though USAREC does recruit from U.S. territories and protectorates, as well as from within American military communities based overseas, this study addresses only recruiting efforts and production in the fifty states plus the District of Colombia.

The data used for this thesis was compiled by the Defense Manpower Data Center (DMDC) for the Navy College Fund Evaluation Study. The period examined is from July 1992 to September 1997. The advantage of analyzing this period was that it is a time of great change, which offers the potential to provide greater contrast in certain indicators that can be exploited. The disadvantage is that it does not address current resources levels and economic conditions.

During the period analyzed by this study, USAREC underwent two major organizational changes with respect to its subordinate units, which are called brigades. Initially, it maintained a five-brigade structure. During 1992, it changed to a four-brigade structure, and then returned to five brigades in 1995. This study assumes that the structure of the brigades was constant, using the current five-brigade organization and its associated geographical boundaries.

C. RESEARCH OBJECTIVES

A considerable amount of past research has been dedicated to the topic of Army recruiting. One of the goals of this thesis is to validate factors from previous models on more current data. Many observers have proposed new or changing influences on the recruiting environment. A further objective of this thesis is to explore these suppositions quantitatively by combining new factors with ones previously shown to be significant. Enumeration of the differences in the recruiting environment throughout the country is another objective. Finally, this thesis aims to develop an accurate tool for predicting recruiting production that can be used by Army leaders.

D. ORGANIZATION

This introduction provides the objectives and organization of this thesis. A detailed overview of Army recruiting and a review of previous research on this subject is contained in Chapter II. Chapter III describes the factors in the time series models and the motivation for their inclusion. The modeling methodology is developed in Chapter IV. Chapter V contains the results and a discussion of their implications. Finally, conclusions and recommendations are provided in Chapter VI.

II. ARMY RECRUITING

A. OVERVIEW

1. Mission and Structure

The Army currently requires approximately 75,000 new soldiers a year. The Office of the Deputy Chief of Staff for Personnel (ODCSPER) determines this requirement and passes it to the U.S Army Recruiting Command, the organization responsible for recruiting civilians for service in the Army. Recruiting Command is organized into five subordinate brigades, which have general regional responsibilities as follows: northeast, southeast, north central, south central, and west. In many cases, states are divided between different regions. The current organization's geographic boundaries are reflected in figure 2.1.

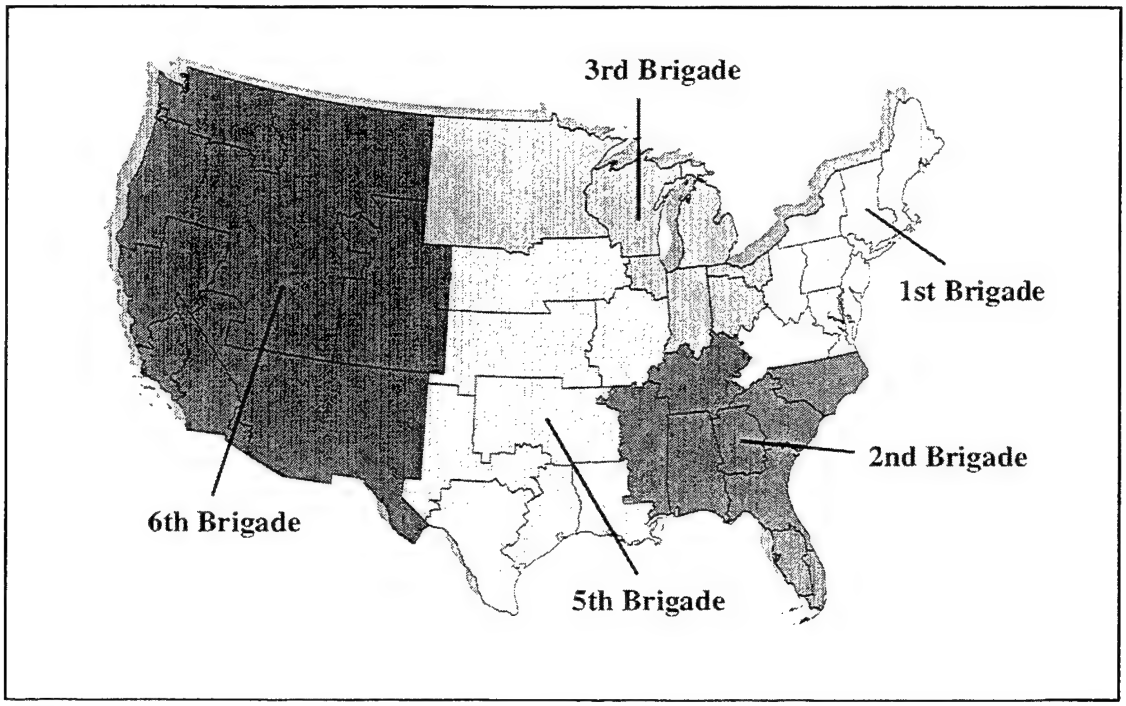


Figure 2.1 U.S. Army Recruiting Command Structure

USAREC has approximately 1,600 recruiting stations where the business of making contacts and signing contracts actually takes place. In addition to its 6,000

recruiters, USAREC currently employs an additional 4,250 uniformed personnel and 1,100 civilians in research and support activities. Beyond people, one of its major resources is advertising. In 1997, USAREC spent approximately $87.9 million in national television, radio, and print advertising campaigns. This figure does not include additional funding that is provided to subordinate commanders for local advertising efforts.

2. High-quality Recruit Definition

The Army seeks to fill its ranks with "high-quality recruits." Such people are trainable on technically oriented jobs. They also have higher contract completion rates, and greater retention for additional contracts. The Army defines a high-quality recruit as one who is in Test Score Category (TSC) I-IIIA, meaning he or she scored above the 50th percentile on the Armed Forces Qualification Test. Additionally, a high-quality recruit is in Educational Credential Tier 1, which means that he or she is a high school or GED diploma holder. The Department of Defense (DOD) and the Army state policy objectives for the number of TSC I-IIIA individuals recruited. U.S. law, DOD, and the Army have increasing minimum requirements for high school degree holders respectively. The Army policy dictating Tier 1 accessions has varied between 90 and 95% of all accessions over the past decade.

3. Recruiting in the 1990s

The 1990s represented a period of great change for the United States Army. The decade began with the Cold War victory and was followed in 1991 with the victory in the Gulf War. Since then, the Army has experienced increasing operational tempo with numerous peacekeeping and humanitarian deployments to Somalia, Haiti, Bosnia, and Kosovo, among others. Correspondingly, the 1990s represented a turbulent period for the Recruiting Command.

The post-Cold War drawdown reduced the Army's strength approximately 35%. From 1989 to 1998 accession requirements decreased by about the same degree (Asch, 1999). The combination of a reduced demand for new soldiers, the military's increased public popularity following the triumph in the Persian Gulf, and the 1992 economic recession, allowed the Army to recruit the best educated force in its history. In 1991, 98%

of new soldiers were high school graduates (Eitelberg, 1994). Under these conditions, Recruiting Command was able to cut recruiter strength 25% and reduce advertising budgets over 50% between 1989 and 1994 (Orvis, 1996). USAREC also reduced overhead in the organization of its subordinate commands. In 1992, it consolidated its subordinate units from a five-brigade to a four-brigade structure.

Towards the conclusion of the drawdown, the Army entered a "steady state," meaning that every soldier who left the service had to be replaced by a new recruit. As a result, accession requirements began to increase slightly in 1995. In response, USAREC returned to a five-brigade structure, and has increased recruiter strength 15% since 1997. It is now initiating a Corporal Recruiter Program to employ younger soldiers to better relate to its target audience (Dickey, 1999). USAREC is also offering shorter enlistment terms, and in 1999 began, for the first time, to combine enlistment bonuses with the Army College Fund. Despite these efforts, the Army has failed to meet its recruiting requirements every year since 1997.

Several factors in the recruiting environment have contributed to the Army's recent shortfalls. The lack of a military threat to the nation decreases the perceived need to serve. By 1999, the military-to-civilian pay gap had grown to 13.5%, its widest level since 1979 (Parlier, 1999). The country has experienced the lowest sustained unemployment rate in thirty years (Bureau of Labor Statistics, 1999). Finally, the increasing number of youth attending post-secondary education and the increasing financial return on a college degree are two related trends that are thought to have had a significant impact on the market for high-quality youth. In just a four-year period starting in 1990, the number of 18-19 year-old youths attending post-secondary education increased 5% to 60.2%, and the number of 20-21 year-olds increased 13% to 44.9% (Asch, 1999). Despite an increase in supply of college graduates, the wages they earn have continued to increase relative to those of high school graduates. This indicates that the demand for the skills that these graduates bring to the workplace continues to be greater than the supply.

B. REVIEW OF PREVIOUS RECRUITING RESEARCH

A significant number of studies focus on predicting recruiting production for the All-Volunteer Force. This research provides insight into influential factors and the methods used to identify them. The following two studies do not include results from multivariate regression, but do provide useful background information on important variables and trends in Army recruiting.

1. General Background Studies

Orvis, Sastry, and McDonald's 1996 *Military Recruiting Outlook* breaks the recruiting process into 2 major factors: "supply of potential enlistees" and "conversion of potential supply" (Orvis, 1996). The researchers employ single-variable regression of specific indicators to identify trends in propensity and in conversion of potential supply. The authors determine that the predicted supply for FY 94 and 95, as measured by propensity of high-quality recruits to enlist from the YATS results, was actually greater than pre-drawdown levels of supply. This suggests that any shortfalls in recruiting for these two years resulted from the Army's inability to convert supply to enlistments. The study reveals that the trend in propensity to enlist was decreasing, especially for minorities. The authors predict (accurately in retrospect) that by FY 97 this trend, combined with the increasing post-drawdown accession requirements, would result in the service facing a supply shortage in addition to its conversion difficulties. The study recommends further research to identify causal factors for the conversion shortcomings.

Asch, Kilburn, and Klerman's 1999 RAND study, *Attracting College-Bound Youth into the Military*, suggests that recent recruiting shortcomings are a result of permanent changes in the civilian labor market. Specifically, they state that the increase in the college premium, which is the difference between the average real wage of a college degree holder and that of a high school diploma holder, is driving more high-quality youth to seek post-secondary education. Hence, their research indicates that all services are increasingly competing against higher education and not the immediate labor market for TSC I-IIIA recruits. The researchers use existing economic models of

recruiting supply and conduct statistical analysis of various factors to arrive at their policy recommendations.

2. Multivariate Time Series Studies

The following five studies use multivariate regression analysis and/or time series regression analysis to examine factors effecting recruiting production. All five focus on the high-quality enlistee category.

Robert Cotterman wrote *Forecasting Enlistment Supply: A Time Series of Cross Sections Model* for a 1986 RAND study. The author develops a model that predicts monthly enlistment rates for each service in each state based on three empirical factors and 68 indicator variables. Cotterman uses monthly state-level data for each service over a 78-month period starting in 1974. One of the model's distinguishing features is that the covariance structure allowed correlation in disturbances across periods, across services, and across- and within- state components. By using a time series of cross-sections the author avoids collinearity problems associated with using purely time-series data (Cotterman, 1986). The first factor in the model represents the position in the business cycle by a measure of a state unemployment rate's deviation from its trend. The second factor is a ratio of military compensation to manufacturing wages. The last empirical factor is a ratio of recruiting force strength to the target male population size. Indicator variables include month, state, and GI Bill availability. The model's forecasts for FY 81 differ from the actual results by 2% to 13%. It is most accurate for the Air Force. All predictors demonstrate expected behavior and unemployment is the most significant factor. The author concludes that the covariance structure developed in this model reduces the standard error of the estimates from those in earlier models.

Lewis (1987) constructs a time-series of cross sections regression model of 30 environmental factors on Army recruiting production for TSC I-IIIA males. Lewis groups the factors into five major categories: economic, socio-demographic, recruiting resources, enlistment policies, and enlistment competition. The data used covers the period from FY80 to FY84 and is geographically based on 55 of the existing 56 recruiting battalions. The research concludes that the four most positive environmental factors are relative

military pay, unemployment, recruiter strength, and advertising. The most negative factors are minority representation in the population, college degree density, and the introduction of a less robust college fund program.

Dertouzos and Polich's 1989 RAND study, *Recruiting Effects of Army Advertising,* is one of the first research projects to differentiate between various advertising media. Their study uses monthly data for a three-year period from 1981 to 1984 for 66 geographical areas defined by the boundaries of the Military Entrance Processing Stations (MEPS). The model controls for economic and demographic conditions and intensity of recruiter effort. The dependent variable is high-quality enlistments predicted by the number of low-quality recruits, local supply factors (unemployment rate, manufacturing wages, recruiter strength, bonus programs), advertising intensity by medium, and recruiter activity. The most significant supply factors are recruiter strength and unemployment rate. The researchers compare the marginal return on advertising, recruiter staffing, and cash bonuses and conclude that advertising is the most cost-effective of these three resources. The study reveals that national magazines and local newspapers are the most effective media followed by national radio and network television. Dertouzos and Polich determine that the most cost-effective media are national magazine and newspaper advertisements.

John Warner and Beth Asch summarize the results of a number of empirical models of enlistment supply in their 1995 paper, *The Economics of Military Manpower.* They state that there have been two generations of models since the beginning of the All-Volunteer Force. The general form of the first models is $\ln H = \beta \ln X$, in which H represents the number of high-quality enlistees and X represents a vector of supply variables. The advantage of the logarithmic form is that the variable coefficients could be easily interpreted as "supply elasticities" (Warner, 1995). The authors declare that the second generation of models first appeared in 1986 and began to account for the behavior of recruiters. The form of these modes is $\ln H = \lambda \ln L + \beta \ln X + \ln E$, where H and X represent the same elements as the earlier models. L represented the number of low-quality recruits and E represented a measure of recruiter effort based on quotas. The

results of the second-generation models consistently indicate that unemployment rates and relative civilian-to-military pay ratios are significant factors in the recruiting process. The authors conclude that the number of recruiters is the most significant recruiting resource factor.

Dan Goldhaber published a critical review of the Navy's Enlisted Goaling Model in a 1999 report for The Center For Naval Analyses. The Navy Recruiting Command uses this model to predict high-quality recruit contracts on a quarterly basis. The Navy and Army definition of high-quality enlistees is the same. The model's independent variables include recruiter strength, seasonally adjusted unemployment rates, a military-to-civilian pay ratio, YATS propensity figures, combined Army and Navy advertising expenditures, veteran population figures, and additional indicator variables for demographics, seasonality, and policy measures. In the model, all non-binary variables are in logarithmic form. The model uses an autoregressive form to account for correlation between recruiting production in successive quarters. The model's predictions from 1994 to 1999 are within 10 percent of actual production results. Goldhaber uses data from 1992 to 1998 to analyze the structure and components of the model. He concludes that collinearity of the predictive variables did not cause bias in the predictions. He finds that the existing first-order autoregressive form of the model is appropriate. Finally, Goldhaber suggests that feedback from recruiting success influences advertising budgeting. Hence, he recommends removing advertising as a predictive variable to prevent potential biases in the coefficient estimates and the model predictions.

3. Summary

These studies provide insight into what factors have been influential in predicting recruiting production in the past. Over the period encompassed by these works, the mission and composition of the Army has changed dramatically. The quality of the force as measured by the number of high school graduates enlisting has drastically improved, increasing from 16% in 1979 to over 90% throughout the 1990s. Despite these major changes, in all but one multivariate regression analysis, unemployment is the most influential predictor of recruiting production for high quality enlistees. In that one study,

unemployment ranks second to recruiter strength as the most significant indicator. More recent works suggest that competition with post-secondary education and not with the unskilled labor market is a factor with growing importance in predicting recruiting for high-quality youth.

C. EXISTING USAREC PRODUCTION MODELS

One of the forecasting tools USAREC uses is the Command Level Mission Model (CLEMM). It is a model that predicts production at the Battalion level as a function of major demographic indicators and recruiter intensity. In this model, recruiter intensity is measured by recruiter strength and operational policies (*e.g.* the number of recruiting workdays in a month, which can be controlled by varying the number of mandated working Saturdays). Historically, the model has had an accuracy rate within 5% for the TSC I-IIIA category, but is extremely labor-intensive to support. Though the model is still maintained by USAREC and used by the Enlisted Accessions Branch of the ODCSPER, USAREC has abandoned CLEMM in favor of a predictive model based on recent production performance (Pettit, 1999). Beyond CLEMM there are no large-scale models currently in use by USAREC that predict production by incorporating policy, resource, demographic, and economic predictors (Kaylor, 1999).

III. DATA

A. FACTOR SELECTION AND DESCRIPTION

Sixteen factors are initially included for examination in this study as predictive variables. They are selected based on their appearance in previous models or in recent research. The intent of selecting this large number of factors is to determine if factors included in older models are still significant indicators for predicting recruiting production and to determine if new factors postulated to be influential are in fact significant. Unless otherwise specified, all data was provided by the Defense Manpower Data Center (DMDC). This data was originally compiled for a study of the Navy College Fund being conducted by Dr. John Warner of Clemson University. The factors are described below.

The first factor is mission. It reflects the numerical goal for male high school graduates and high school senior in AFQT categories I-IIIA (high-quality) contracts set by Recruiting Command Headquarters for its subordinate Recruiting Brigades to meet each month. It is selected to account for the effort that the production recruiters and their commanders expend in order to meet their assigned recruiting mission. It is also selected to implicitly account for incentives, awards, and bonuses offered to the recruiters on the assumption that the magnitude of rewards are adjusted to correspond with the demands of the mission.

The second factor selected is recruiter strength, which reflects the number of production recruiters assigned to each brigade. Production recruiters are the noncommissioned officers whose job is to make contacts with potential recruits and write enlistment contracts. Production recruiters represent a critical subgroup of the personnel assigned to Recruiting Command and in this context are distinct from commanders, staffs, and government civilians. In previous studies, recruiter strength has been identified as one of the most cost-effective factors effecting recruiting production.

The third, fourth, fifth, and sixth factors represent the impressions made by Army advertising in television, radio, magazines, and newspapers respectively. These factors measure the total audience exposures for each national advertising campaign in a month. They do not reflect the impact of local area marketing programs. The television impressions are measured for advertising run on both network and cable programming. The newspaper impressions combine papers with national distribution along with college campus newspapers. The Army's contracted advertising agencies provide impression figures for various demographic groups. For this study, the audience addressed is 15-24 year-old males.

The percent of eligible recruits receiving the Army College Fund (ACF) when enlisting is the seventh factor. The ACF provides an incentive for youth to join the Army with the promise of dedicated money for post-secondary education upon completion of service. The ACF is only available to potential recruits in TSC I-IIIA. It offers funding in addition to the Montgomery G.I. Bill, which is offered to all Tier 1 enlistees. Research conducted by Beth Asch in her 1999 RAND study suggests that as a greater proportion of young people attend college, this program may be increasingly effective in attracting college-bound youth to enlist.

The eighth factor selected is the target population size. This figure represents the total males in AFQT categories I-IIIA in each recruiting region. This factor is included because Eitelberg and Mehay (1994) predict that a decreasing youth population will compound recruiting challenges.

The unemployment rate is the ninth factor selected. This figure is the ratio of unemployed to the civilian labor force expressed as a percentage. The data is directly extracted from the Bureau of Labor Statistics' Local Area Unemployment Statistics (BLS, Selective Data Access). The unemployment rate chosen is not seasonally adjusted, because the models developed in this thesis include indicator variables for month. Unemployment appears as a significant factor in all models reviewed. It represents the competition for youth that the service faces from the civilian labor market.

The next factor is high school graduate wages, a measure of the average weekly wage earned by male high school graduates in each state. This data is extracted from the Monthly Current Population Survey, which is a joint project of the Bureau of Labor Statistics and the Census Bureau. Individuals surveyed are males age 18-35. To be included in the survey, an individual must have normal weekly hours of at least 30 greater hours. Like unemployment, youth wages are included in all of the previous models examined.

The overall 17-21 year-old college attendance rate represents the eleventh factor. This rate is determined by dividing the college population for each state by the total youth population. These figures are extracted from data compiled by Woods and Poole Economics, Inc., an independent firm that produces county-level economic and demographic projections. DMDC provided this data for use in this thesis. The attendance rate in this model is not specific to gender or to the Army's "high-quality" criterion. This factor is another measure of the competition for bright young people between the military and post-secondary education.

The college premium represents the final factor addressing the Army's competition with colleges for qualified personnel. In this thesis, the figure represents the difference in weekly wages between male high school graduates and college graduates. Like the tenth factor, it is derived from the Monthly Current Population Survey.

The thirteenth, fourteenth, and fifteenth factors represent the monthly recruiting success, measured in signed contracts, of the Air Force, Marine Corps, and Navy in the high-quality male demographic. This category is included to determine if the relationship between the recruiting efforts of other services is competitive or complementary.

The final eleven factors are binary indicator variables for each of the months from February through December. January represents the baseline month and is not specifically represented by an indicator variable.

B. DATA AGGREGATION

The original data for the response variable and the fifteen non-month predictor variables was in various forms regarding their geographic and time divisions. The following table reflects the original form of each variable:

Variable	Index	Geographic Division	Time Division
Army high-quality male contracts	NA	County	Monthly
Recruiting mission	1	County	Monthly
Recruiter strength	2	County	Monthly
TV advertising impressions	3	State	Monthly
Radio advertising impressions	4	State	Monthly
Magazine advertising impressions	5	State	Monthly
Newspaper advertising impressions	6	State	Monthly
Percent of eligible recruits receiving the college option	7	State	Monthly
Target population size	8	County	Yearly
Unemployment	9	State	Monthly
High school graduate wages	10	State	Yearly
College attendance rate	11	State	Yearly
College wage premium	12	State	Yearly
Air Force high-quality male contracts	13	County	Monthly
Marine Corps high-quality male contracts	14	County	Monthly
Navy high-quality male contracts	15	County	Monthly

Table 3.1 The Original Form of the Data for the Selected Variables.

In order to develop a separate model for each recruiting brigade, the original data is aggregated geographically to reflect USAREC's current regional boundaries. The

response variable and six of the independent variables are enumerated at the county level in the data provided by DMDC. A brigade-to-FIPS county code file provided by USAREC is used to index the data for each of the 3,116 individual counties to its appropriate recruiting brigade. The data is then summed for each brigade for each month. For each of these seven variables there are 215,004 records (3,116 counties multiplied by the number of months in the series). Due to the large size of these files, the aggregation procedure is executed using SAS programs developed by Dennis Mar of the Naval Postgraduate School Systems Management Department. For each variable, the number of cases in which records have a FIPS identifier that does not exist in the indexing code is less than 0.10%.

The data for the variables originally organized at the state level is converted to the appropriate regional structure using weights derived from the target population. The weights are calculated using

$$weight_{sr} = \frac{targetPopulation_{sr}}{targetPopulation_{s}},$$

where the subscript *sr* indicates the state, *s*, in region *r*. For states that are divided among regions, the numerator for the weight calculation is the portion of the state's target population in each region. Once the weights are calculated, they are multiplied by the state figure for an independent variable. These values are then summed over states to determine a figure for the brigade, as shown:

$$indepVariable_{r} = \sum_{s} (weight_{sr} \cdot indepVariable_{s}).$$

The four independent variables originally represented in annual time divisions are converted to monthly figures by developing a linear relationship between the data points. Monthly figures are determined by

$$indepVariable_{ym} = indepVariable_{y} + \left(\frac{indepVariable_{y+1} - indepVariable_{y}}{12} \right) \cdot monthNumber$$

where y represents year and m month. Figures for target population size and college attendance rates are calculated assuming that the original data points are for the month of January. The figures for high school graduate wages and college premium are calculated

assuming that the original data points are for the month of June. The *monthNumber* factor reflects the months, ordered 1-12, between the original annual data points. This linear transformation of annual observations has the potential to produce additional noise in the process. Original data with monthly observations is preferable, but is unavailable. However, target population figures and college attendance rates are not expected to change significantly month-to-month. The greatest potential for error induction is for the college premium and especially the high school graduate wage level, both of which could potentially exhibit seasonal behavior.

C. EXPERIMENTAL DESIGN AND DATA SEGREGATION

The original data represents a 63-month period from July 1992 through September 1997. The first 54 months are selected as an analysis data set for training the models. For all model development, *T*, the number of observations a series, is initially equal to 54. The remaining nine months are reserved as the validation data set to test the accuracy of the models' forecasts. The series extrema and averages for the full, test and training data sets are listed in Appendix A.

IV. METHODOLOGY

A time series is a sequence of observations made over time. The guiding principle behind time series analysis and forecasting is that the future can be predicted based on determination of patterns in past data (Bowerman, 1979). In multivariate time series analysis the analysis task is expanded to include the determination of the interrelationship between multiple series (Chatfield, 1996). Based on the relatively small size of the recruiting data set and the varying scales of the regressor values, determining which factors were significant is problematic. The means employed to overcome this difficulty is the bootstrap (Efron, 1998). It consists of resampling from within the existing data to provide robustness in the factor determination process. Use of the bootstrap technique requires that the residuals of a hypothesized model be independent. This necessitates the development of an autoregressive moving average (ARMA) model for each brigade.

These requirements lead to the following methodology. First, select an appropriate time series model to produce residuals with structure suitable for bootstrapping. Second, develop the bootstrap recursion. Third, develop a recursion to conduct stepwise reduction of the model to identify significant factors. Fourth, develop and perform diagnostics on a final reduced model. Finally, use this model to predict future recruiting production and determine the accuracy of the predictions. The process described in this chapter addresses the steps to develop one model and one nine-month forecast. It is applied five separate times to develop a model and a forecast for each USAREC brigade.

A. FITTING MUTIVARIATE ARMA MODELS

For most measurements taken at fixed intervals over time, there is an underlying structure to the data. That is, there exists an association between one observation and its "neighbors." One of the primary tasks of the analysis is to determine the strength of that relationship. A multivariate time series also accounts for the influence of regressor series

on the response variable. Models of the following form capture these relationships for the recruiting data:

$$\begin{aligned} y_t - \mu_y = \beta_1(Mission_t - \mu_1) + \ldots + \beta_{27}December + \phi_1(y_{t-1} - \mu_y) + \ldots \\ + \phi_p(y_{t-p} - \mu_y) + \varepsilon_t + \theta_1\varepsilon_{t-1} + \ldots + \theta_q\varepsilon_{t-q} \ , \end{aligned} \qquad \textbf{(3.1)}$$

where y_t is the response variable, high quality male contracts, at time t. The β_i specify the relationship of the value of the regressors to the response variable. For both the response and the non-binary predictive variables the data must be centered on μ, the mean value of the respective series. The autoregressive parameters, ϕ_p, capture the strength of the relationship between the value of the response variable at time t and observations in previous periods. Uncertainty in the time series process is captured by ε_t, a "purely random disturbance term with a mean of zero and a variance of σ^2" (Harvey, 1994). The moving average parameters, θ_q, represent the relationship between the response and these disturbances in previous periods. The number of autoregressive parameters, p, and the number of moving average parameters, q, define the order of an ARMA (p, q) model. In essence, the multivariate ARMA model for each brigade represents the deviation of the response variable at time t from the mean of its series by the deviation of the non-binary predictive variables from their respective series means, the magnitude of the monthly effect, plus the autoregressive and moving average effects.

Equation 3.1 models the complete series of recruiting data. The values of the model parameters are estimated through analysis from a sample of this theoretically infinite series. A number of notational changes are made to distinguish these estimations from the parameters of the complete series. The observed values of the response variable and the predictive variables, μ_y and μ_i, are represented by $\overline{y}_i$ and $\overline{x}_i$, where the index $i = 1,2,\ldots,15$ runs over the indices listed in Table 3.1. The actual regression coefficients, actual autoregressive and moving average parameters, and disturbances are estimated by $\hat{\beta}_i$, $\hat{\phi}_p$, $\hat{\theta}_q$, and $\hat{\varepsilon}_t$ respectively. To simplify the representation of the terms in equation (3.1) and to address the distinctions of sample data, the following notation is introduced.

The deviation of the response variable, high-quality male recruiting contracts, from the mean of its series is represented by

$$z_t = y_t - \bar{y},$$

while the value of the response variable estimated by a model is represented by $\hat{z}_t$.

The centered regressors are represented by

$$s_{it} = (x_{it} - \bar{x}_i),$$

where the index i = 1,2,...,15 on each respective x_{it} corresponds to the indices listed in Table 3.1. The models developed based on the sample data therefore have the form

$$\hat{z}_t = \hat{\beta}_1 s_{1t} + ... + \hat{\beta}_{27} December + \hat{\phi}_1 z_{t-1} + ... + \hat{\phi}_p z_{t-p} + \hat{\varepsilon}_t + \hat{\theta}_1 \hat{\varepsilon}_{t-1} + ... + \hat{\theta}_q \hat{\varepsilon}_{t-q}. \quad \textbf{(3.2)}$$

Models of the form reflected in equation (3.2) are created using the Gaussian maximum likelihood estimation method in the S-Plus statistical software package (Mathsoft Inc., 1999). Two criteria are used for selecting the appropriate ARMA model for each brigade. First, the model's residuals must not display any significant autocorrelation or partial autocorrelation. Second, the model has to be of the lowest possible order while fitting the data well.

The autocorrelation, ρ_k, represents the strength of the relationship between any two observations in a time series separated by a lag of k time periods. The autocorrelation is a dimensionless measure with values between 1 and –1. When ρk is large in magnitude, observations k time units apart tend to move together in a linear fashion, and hence are not independent. The sign of ρ_k indicates the direction of this movement. The partial autocorrelation, ρ_{kk}, represents the autocorrelation between any two observations separated by a lag of k ignoring the effects of the intervening observations. The autocorrelation function (ACF) and partial autocorrelation function are lists of ρ_k and ρ_{kk} at lags k=1,2,3...(Bowerman, 1979). The graphs of these functions are called correlograms, which are the actual diagnostic tools used to assess the first criterion. An example correlogram with lags up to k=24 is displayed in Figure 3.1.

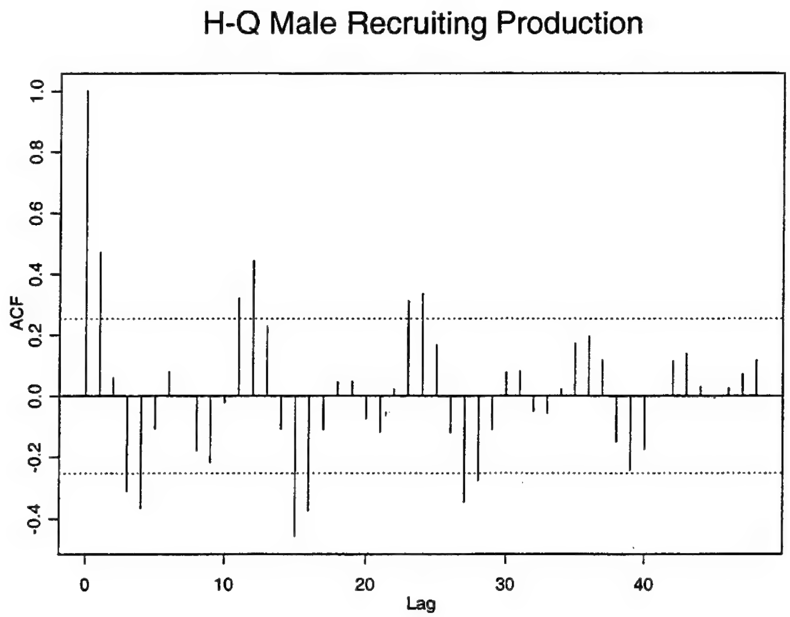


Figure 3.1 Example Correlogram Produced Using S-Plus

The standard for determining significant ACF and partial ACF involves Bartlett's approximation of the standard error of autocorrelation estimates, var[r_k] (see Box and Jenkins, 1976, pp. 34-35). Values of the ACF at any lag $k > 0$ which exceed $+/-2\sqrt{\text{var}[r_k]}$ are considered significant. This range is automatically calculated by S-Plus and indicated by the dotted horizontal lines in the generated correlograms.

Box and Jenkins define the notion of developing parsimonious ARMA (p, q) models, which means choosing the smallest values of p and q that adequately capture the nature of the time series. The Akaike information criterion (AIC) captures the essence of this concept. It provides a tool for comparing models, by indicating a model's goodness of fit while penalizing complexity. Models are selected by minimizing the AIC, which is defined by

$$AIC = -2 * \log L(\psi) + 2n.$$

In this equation, $L(\psi)$ is the maximized value of the likelihood and n is the sum of the ARMA model orders ($p + q$). Complex models have higher values of $L(\psi)$, hence a large

negative first term, but are penalized for having a large value of n. Simple models have smaller values of $L(\psi)$ creating a smaller negative first term, but small values of n. The AIC is a means of balancing these competing characteristics. (Harvey, 1994)

B. BOOTSTRAP

All of the ARMA models developed have random disturbance terms $(\hat{\varepsilon}_1, \hat{\varepsilon}_2, \ldots, \hat{\varepsilon}_T)$. In models that include an AR term, observations from the first p time periods are used to initiate the autoregressive process and no estimates from these periods are derived. Hence, the first disturbance term from an AR model is $\hat{\varepsilon}_{p+1}$. Since the $\hat{\varepsilon}_t$ represent random disturbance terms, by definition they are assumed to be independent.

The ARMA model residuals are defined by

$$e_t = z_t - \hat{z}_t$$

in which e_t is the deviation of the selected model's response variable value from the original observation at time t. Like the random disturbance terms, the residuals are assumed to be independent based on the first modeling criterion. Because of their similar properties, the residuals are used to approximate the random disturbance terms. This concept is the basis for the development of the bootstrap method in this application.

The bootstrap recursion first samples with replacement from the set of residuals ($e_{p+1}, e_{p+2}, \ldots, e_T$) to develop a new set of residuals, e_t*. Next, a new series of the response variable, z_t*, is simulated using the new set of residuals to represent the random variations, so that $z_t^* = \hat{z}_t + e_t^*$. The final step is to refit an ARMA model of the same order as the original, using the original regressor's paired with the simulated z_t* series. The simulated response variable values force slight changes in the estimates of the $\hat{\beta}_i$, $\hat{\phi}_p$, and $\hat{\theta}_q$. At the conclusion of the recursion, the regression coefficients are saved to a vector. The applicable autoregressive and/or moving average parameters are saved to their own vectors.

Concatenation of the output vectors for multiple repetitions of the bootstrap recursion creates a matrix of regression coefficients. Each column of this matrix represents the developed regression coefficients for one factor. Analysis of the central tendency and variance of the figures in each column provides robustness in determining the value of the $\hat{\beta}_i$ and hence, the influence of each factor on the response variable.

The method of generating the approximations to the random disturbance terms described above represents a non-parametric approach. The original distribution of the residuals used to represent the random disturbance terms is preserved in this method. An alternate approach is also developed in which the random variations are generated by sampling from a normal distribution with a mean of zero and a standard deviation equal to the standard deviation of the set of residuals, $(e_{p+1}, e_{p+2}, \ldots, e_T)$. The S-Plus code for executing this bootstrap recursion allows specification of whether to use the parametric or non-parametric approach for sampling from the residuals. The default method is non-parametric.

The underlying theory for the development of this recursion is due to Efron and Tibshirani (1998, chapter 8). See Appendix B for the bootstrap recursion S-Plus code.

C. STEPWISE REDUCTION

The bootstrap recursion provides a tool for overcoming some drawbacks of having a limited number of time series observations from which to determine the relationship between the predictive and the response variables. The next stage of model development addresses the research objective of identifying the significant factors in predicting recruiting production.

The underlying premise for the elimination of factors is as follows. If the mean value of a regression coefficient, $\hat{\beta}_i$, calculated from multiple iterations of the bootstrap is within a defined interval around zero, it can be interpreted as not significantly different from zero (Efron, 1998). If this is the case, the associated factor is not considered influential in predicting the behavior of the dependent variable. Such a factor can be

eliminated from the model, which promotes model simplicity and improves model accuracy by removing noise associated with the discarded factor.

Experimentation with the application of this idea for identifying non-significant factors provides valuable insight. Initially, factors that had means in a defined range around zero were removed all at once. The reduced ARMA models produced by this approach gave unpredictable results. Specifically, the regression coefficients for the remaining factors demonstrated sign changes from the original to the reduced model. This observation led to the development of a stepwise recursion to promote stability as the model is reduced.

The intent of the recursion is to eliminate factors one-by-one until only significant factors remain in the model. The recursion first executes the bootstrap to develop a matrix of $\hat{\beta}_i$. Factors are identified as candidates for elimination if the mean of a column of regression coefficients lie within the range of $0 \pm \alpha \cdot$(standard deviation of the column of regression coefficients). The α term is an input parameter that controls the width of the interval. For each candidate factor, the proportion of the estimated $\hat{\beta}_i$ in the range $0 \pm \alpha \cdot$(standard deviation of the column of regression coefficients) is calculated. The candidate factor that has the highest proportion is eliminated and a new ARMA model of the original order is refit. This recursion is repeated until no factors are eliminated. The final significant factors are then displayed.

The S-Plus code for executing the stepwise regression allows control of the number of iterations of the bootstrap recursion, the method of residual sampling (parametric or non-parametric), and α, the tolerance defining the size of the interval around zero. The accepted standard for repetition of the bootstrap is 1,000 iterations (Efron, 1998). Non-parametric residual sampling is employed and a value of 1 is used for α. The code for the stepwise reduction recursion is in Appendix B.

D. FINAL MODEL DEVELOPMENT AND DIAGNOSTICS

After the significant factors for a brigade are determined, the final reduced ARMA model is produced. It is of the same order as the original full model. Correlograms are plotted to ensure there is no significant ACF or partial ACF of the residuals. The AIC of the reduced model is calculated to ensure it is less than the AIC of the original full model. The regression coefficients of the factors remaining in the final model are then examined to ensure that they have the same sign as they did in the initial full model. The presence of a sign change is not necessarily an indication that the model is invalid, but it prompts scrutiny of the data and the factor reduction process.

E. FORECASTING

Once the final model for each region is determined, it is used to forecast the number of high-quality male contracts in the test period. Like the training data, the test data is first centered on the mean of respective variable from the training series.

A simulation is used to develop a predicted time series of the response variable. The length of the predicted time series is nine months, corresponding to the length of the test data period. The regressor values are from the test data. The simulation uses a parametric approach to the generation of the random errors. The mean of the random errors is zero and the standard deviation is equal to the standard deviation of the residuals from the final model. The simulation is repeated 1,000 times to develop multiple predictions for the nine-period series. The mean and standard deviation of the 1,000 predictions for each month of these simulated series are used to make the forecast. The code for producing the forecasts is contained in Appendix B.

The forecast error for each observation in the 9-period time series is defined by

$$\Delta_t = z_t - \hat{z}_{t(forecast)}.$$

Once calculated, diagnostics of the forecast errors are performed. The Δ_t values are plotted to determine if they appear randomly distributed. Correlograms of the forecast errors are plotted to determine if they exhibit significant ACF or partial ACF.

In order to provide results that can be conveniently and directly compared to the original data, the centering procedure required for ARMA model development is reversed as follows

$$\hat{y}_{t(forecast)} = \hat{z}_{t(forecast)} + \bar{y}.$$

The percent error of the forecasts, defined as

$$percentError = (y_t - \hat{y}_{t(forecast)}) / y_t \cdot 100,$$

is also calculated for each of the nine forecast values. Finally, the actual behavior of response variable, y_t, the forecast value, $\hat{y}_{t(forecast)}$, and the forecast plus and minus one standard deviation of the forecast are plotted to provide a visual tool for interpreting the forecast's accuracy.

THIS PAGE INTENTIONALLY LEFT BLANK

V. RESULTS

Models are developed that include all sixteen factors plus monthly indicators as described in Chapter III. Based on the recommendations made by Goldhaber in his 1999 review of the Navy's Enlisted Goaling Model, models that exclude all advertising data are also developed. The models that do not include advertising impression data are more accurate for forecasting as measured by the percent error of the forecasts. Therefore, all model results described in this chapter refer to models that initially excluded advertising impression data. The behavior of the advertising time series is still examined in the Descriptive Statistics section.

A. DESCRIPTIVE STATISTICS

1. Time Series Graphs

A basic step in the analysis of time series is plotting the data to identify trends, outliers, seasonality, and other cyclic changes (Chatfield, 1996). The series for the variables in each brigade are reflected in Figures 5.1 through 5.5. In all these graphs, the training and test series are plotted as one series. A vertical line between December 1996 and January 1997 indicates the division between these two sets. Note that the linear construct of the four independent variables converted from annual to monthly data, (target population size, college attendance rates, high school graduate wages, and college premium), precludes observation of seasonal behavior.

The first seven time series graphs address eight variables specific to Army recruiting and USAREC policies. The Army's high-quality male recruiting shortages are clearly reflected by the increasing difference between the recruiting mission and production in each brigade. Recruiting production demonstrates clear seasonality with peaks each June. Recruiter strength does not appear seasonal, but assignments increase noticeably in all brigades beginning in early 1997. All advertising media display a decreased number of impressions around the months of June and July. Television and

radio demonstrate increasing trends, while magazine and newspaper impressions lack clear trends. The percent of youth receiving the Army College Fund exhibits no discernable trend or seasonality.

Factors of the recruiting environment are captured in the next five time series graphs. In all brigades, target population figures initially demonstrate a decrease between 0.2% and 2.9%. However, in all but the First Brigade region, there is a net growth in the target population between 1992 and 1997. Relative growth is greatest in Second Brigade, followed closely by Sixth Brigade (10.1% and 9.9% respectively). In all regions, unemployment displays clear seasonality with peaks in January and June and an overall decreasing trend. High school graduate wages demonstrates a steady increase over time. The behavior of college attendance rates differs between brigades. All regions show a dip in attendance in 1994. In the Fifth and Sixth Brigade regions, there is a net decrease in the college attendance rate over the period examined, while the other three regions experience a net increase. None of the attendance rate changes are more than two percentage points. The college premium exhibits a net increase in all brigades, though the behavior varies by region. The relative increase is largest in Third Brigade (43%) and smallest in Second Brigade (10%).

The final time series graphs address the behavior of rival services' recruiting production in the high-quality male demographic. The Air Force, Marine Corps, and Navy all display a decreasing trend until about 1994, after which the mean of each series appears fairly constant. All the rival services demonstrate seasonal summer peaks, though their occurrence seems to vary by one to two months, with the Air Force's peak occurring later in the summer.

THIS PAGE INTENTIONALLY LEFT BLANK

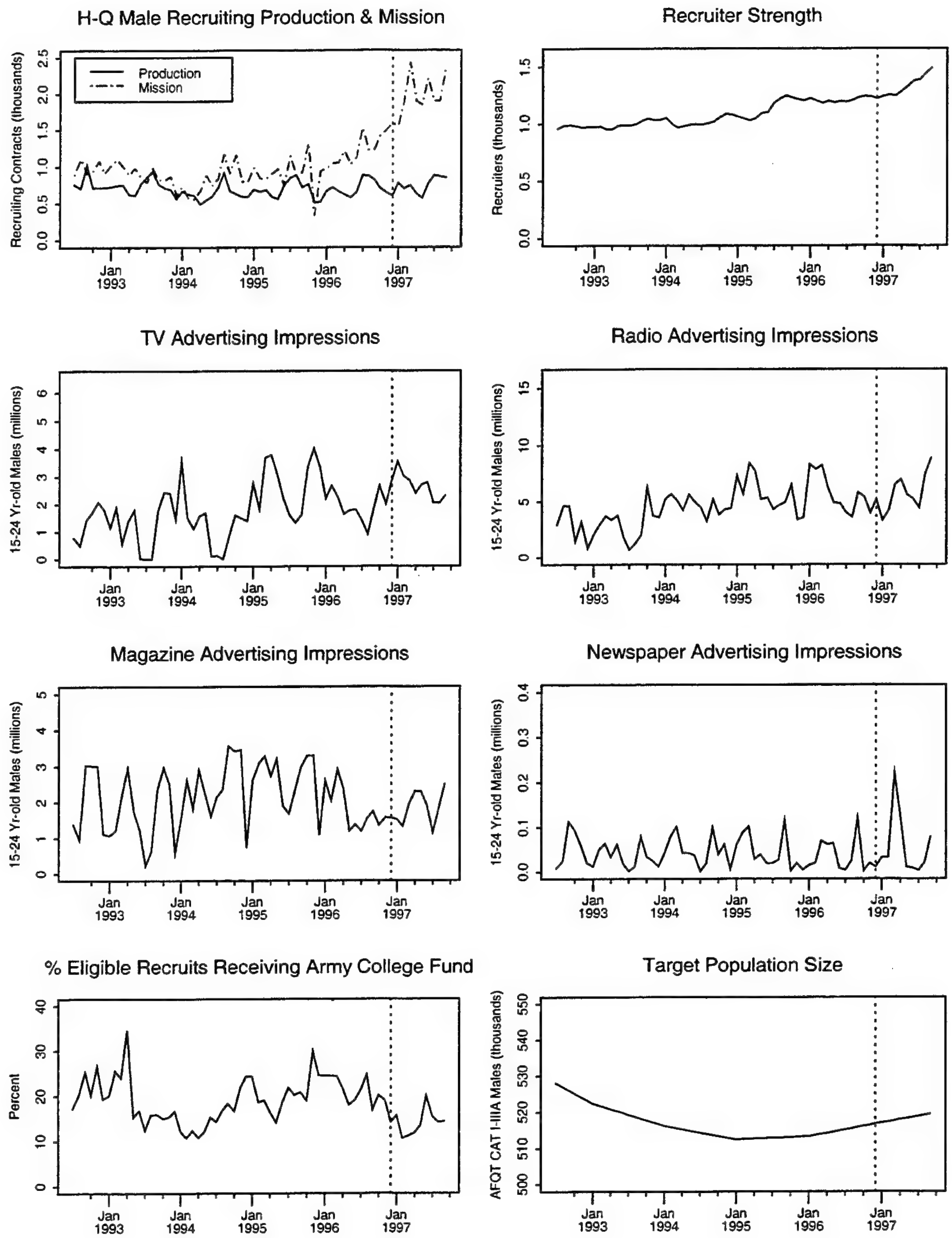


Figure 5.1 First Brigade Factor Time Series

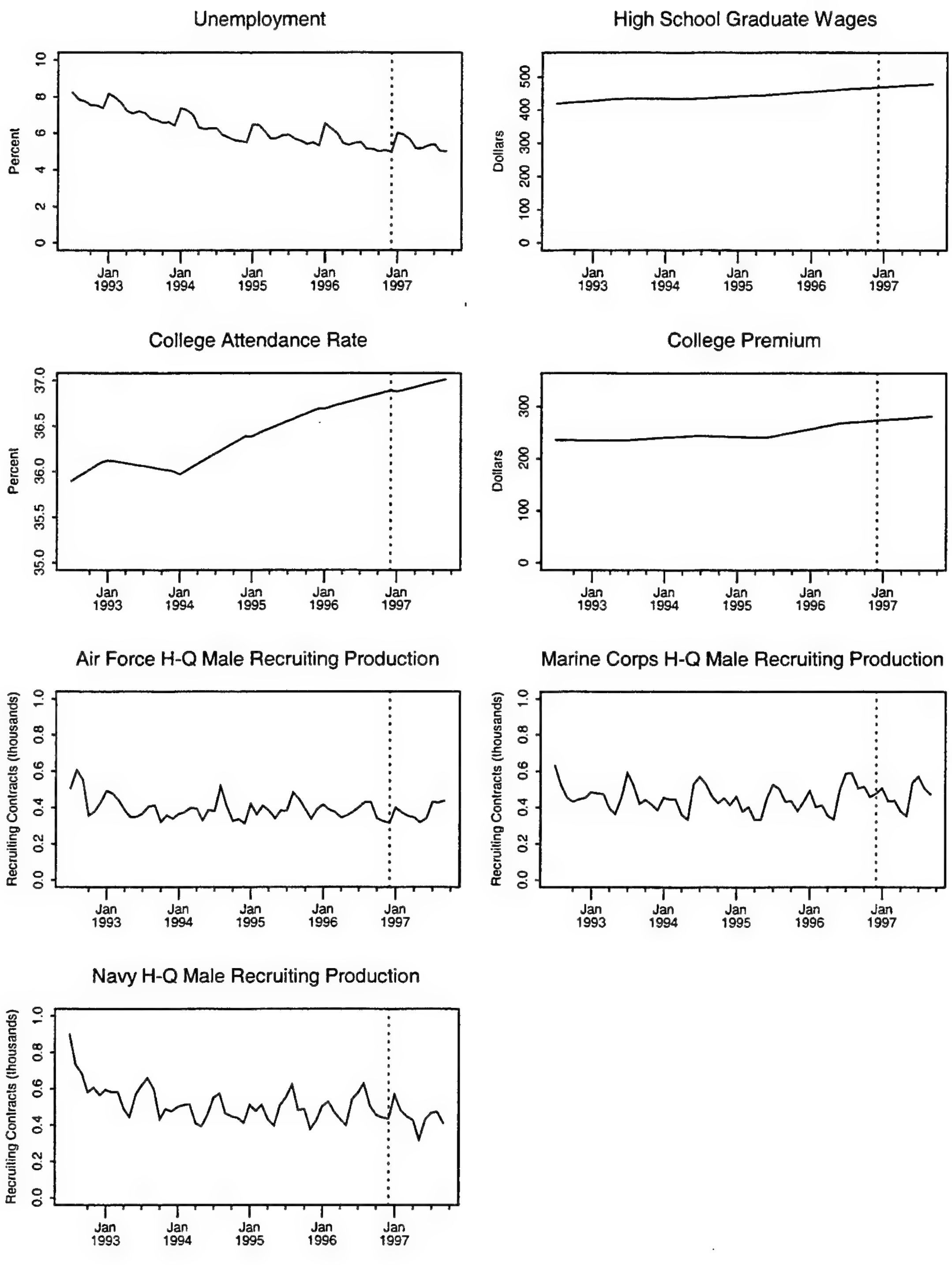


Figure 5.1 Continued First Brigade Factor Time Series

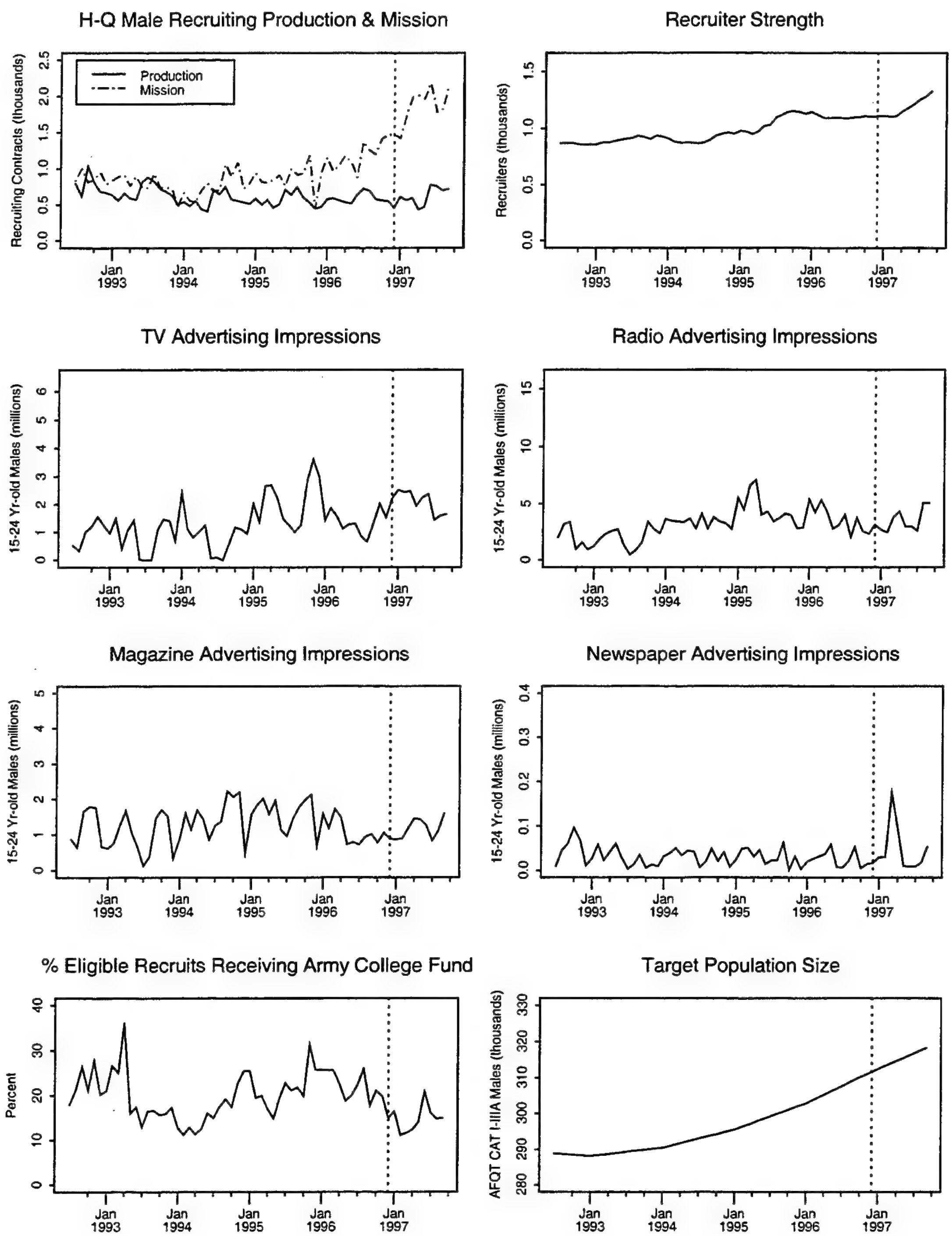


Figure 5.2 Second Brigade Factor Time Series

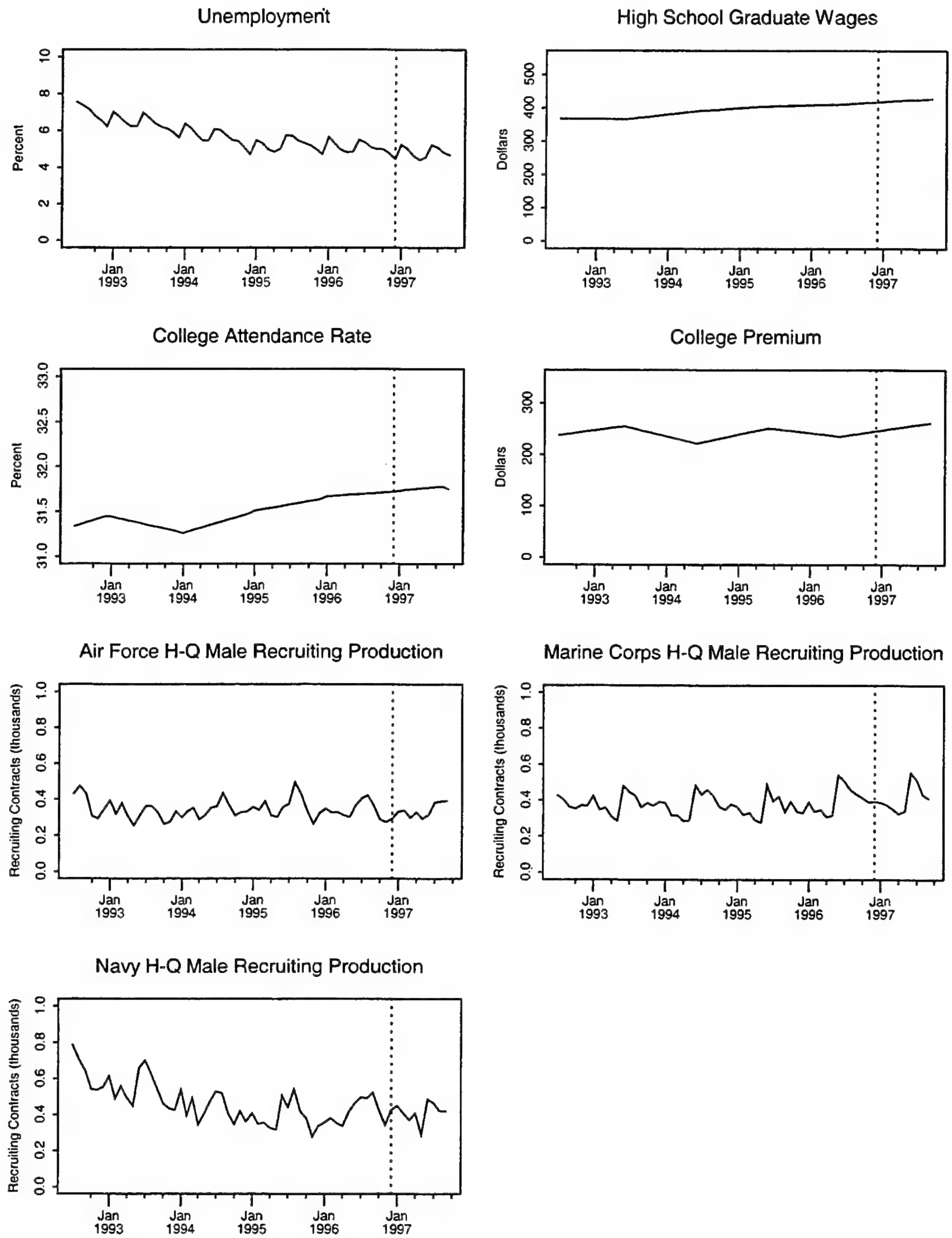


Figure 5.2 Continued Second Brigade Factor Time Series

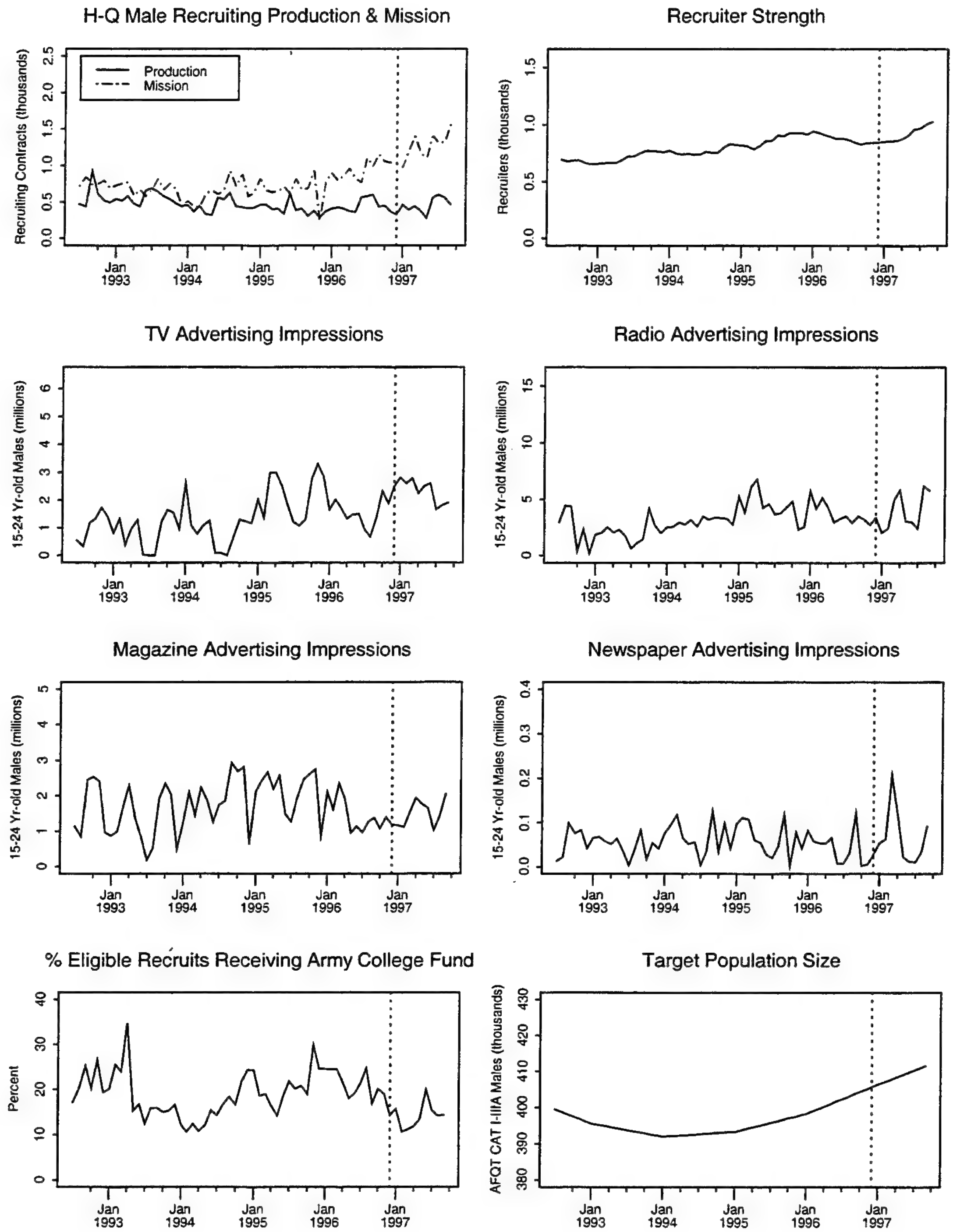


Figure 5.3 Third Brigade Factor Time Series

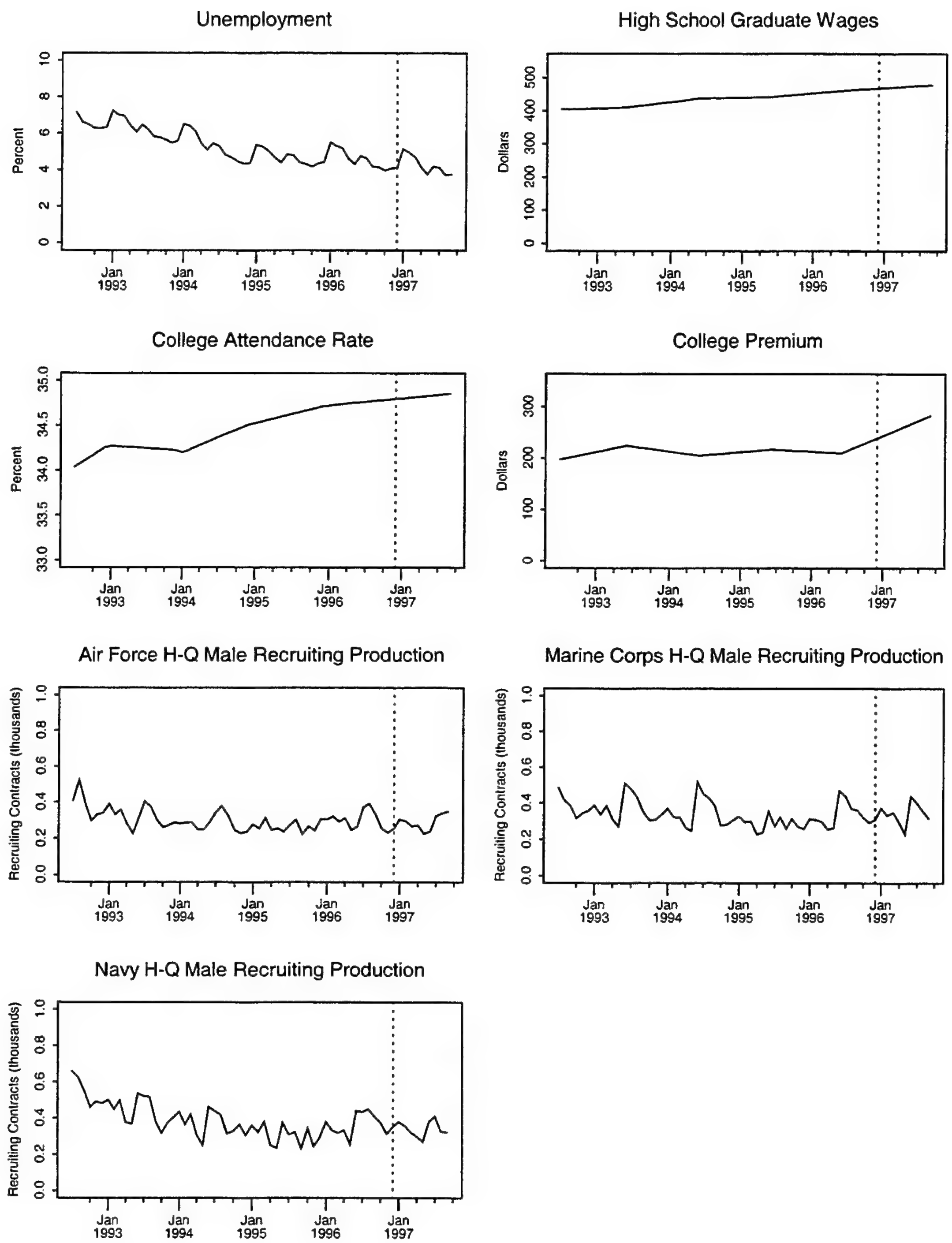


Figure 5.3 Continued Third Brigade Factor Time Series

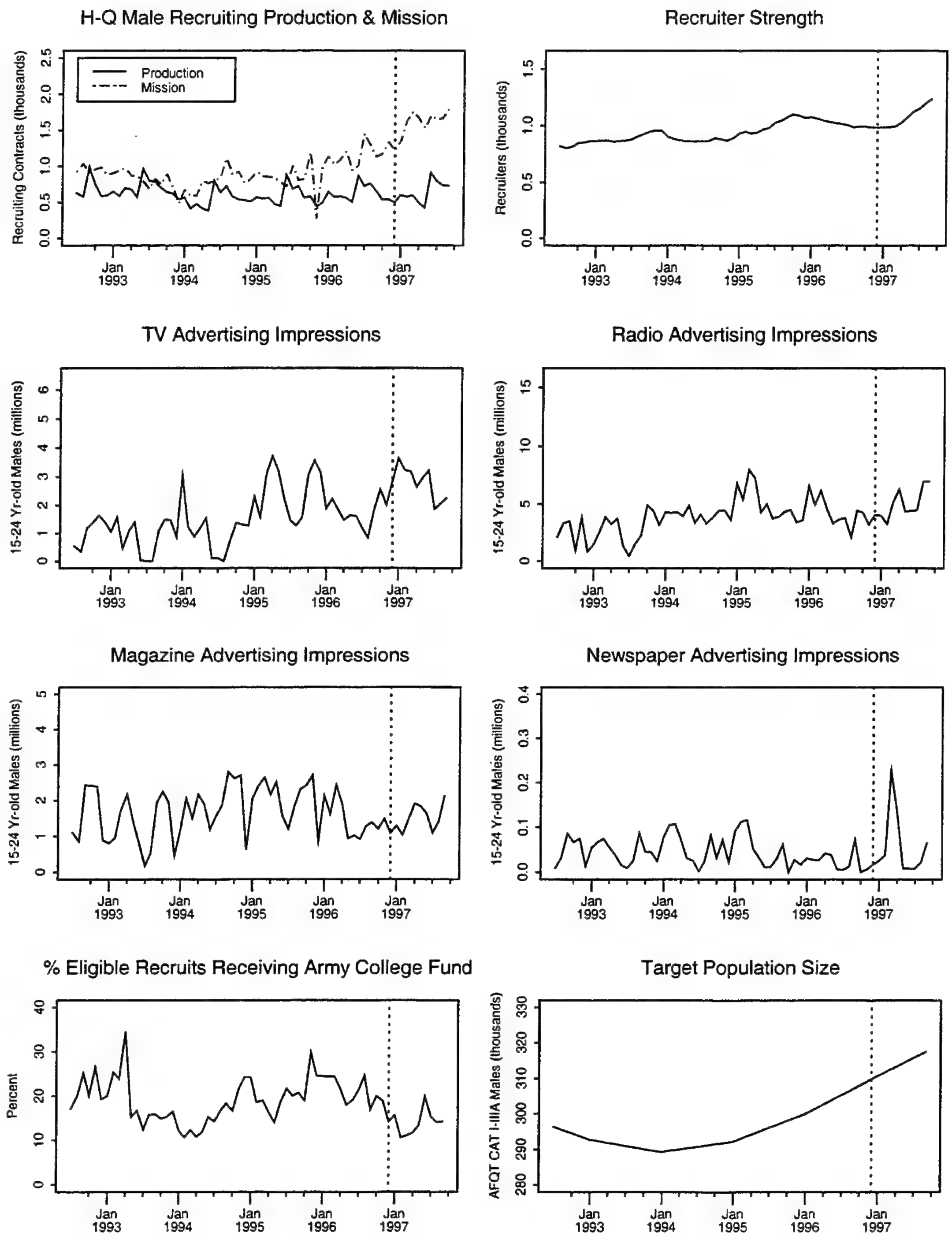


Figure 5.4 Fifth Brigade Factor Time Series

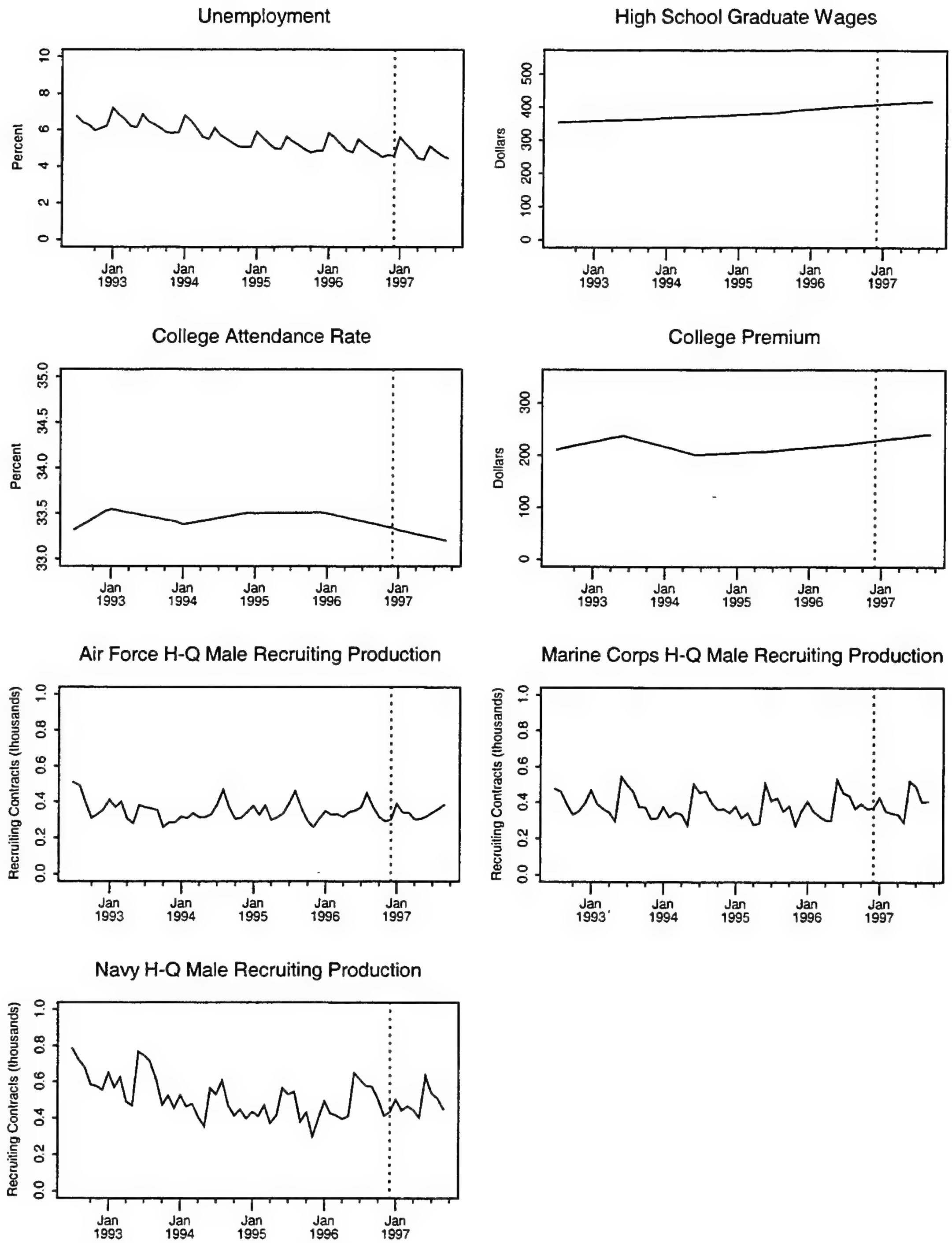


Figure 5.4 Continued Fifth Brigade Factor Time Series

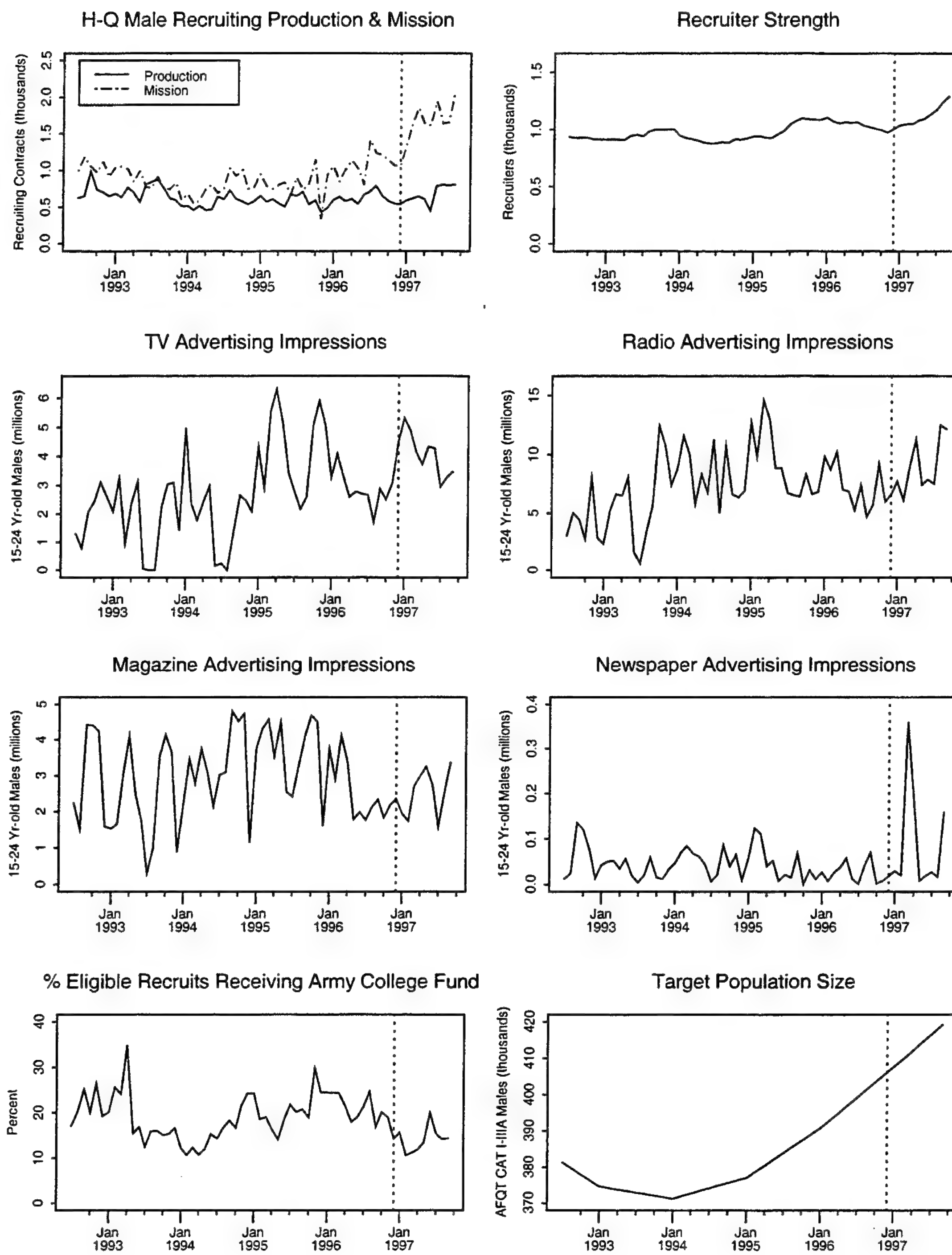


Figure 5.5 Sixth Brigade Factor Time Series

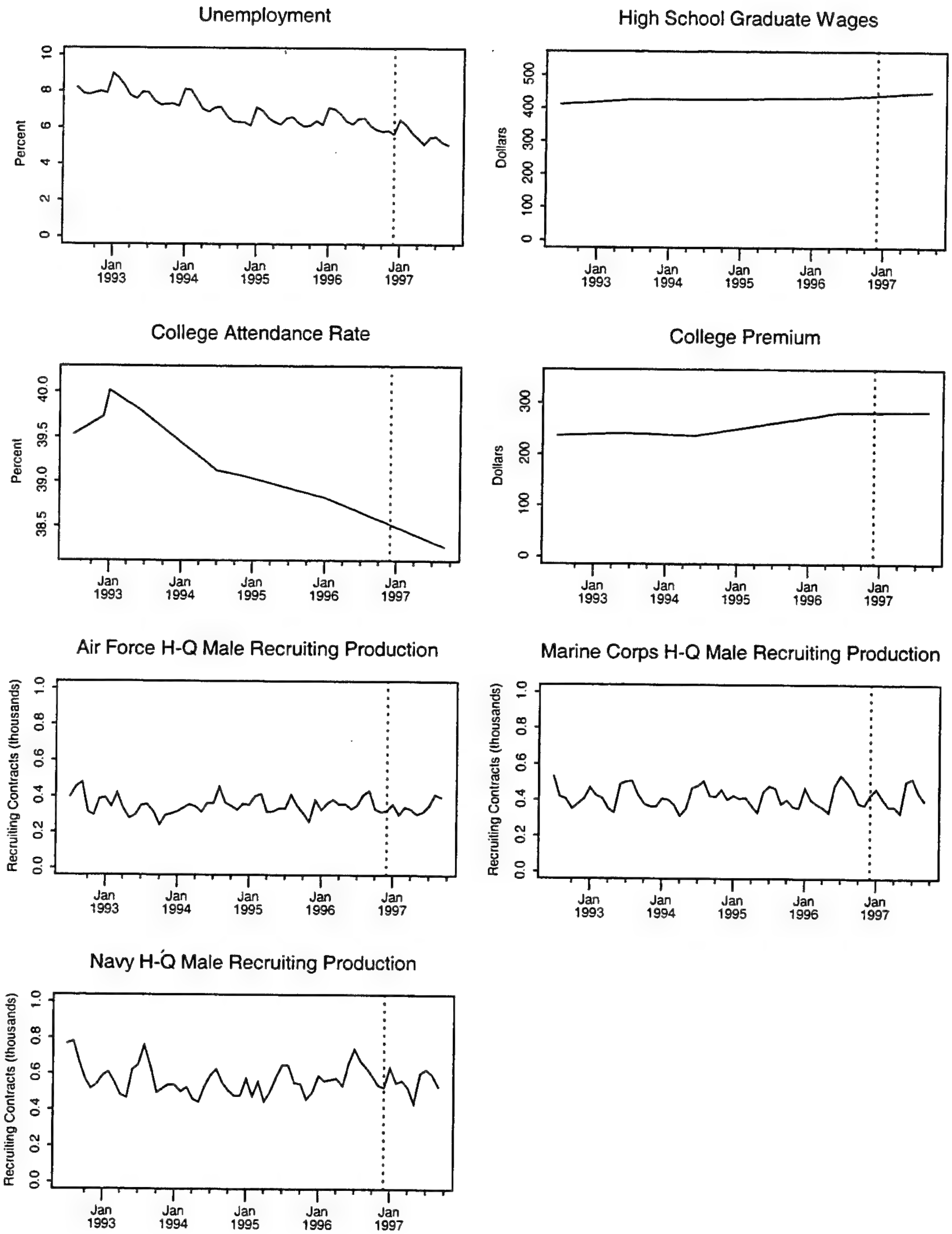


Figure 5.5 Continued Sixth Brigade Factor Time Series

2. Time Series Structure of the Variables

The correlograms for the variable series in all brigades are contained in Appendix C. They prove valuable for confirming the observations of seasonality discerned from the time series graphs and also for identifying cycles that are not visually evident. Only departures from graphical observations, additional insights, and unexpected results are addressed in this section.

The ACF of the mission variable series decays to an insignificant level in a lag of k=5 or less in all brigades. This behavior is somewhat unexpected. It suggests that USAREC did not issue contract missions to its subordinate brigades in accordance with the known seasonal behavior of recruiting production.

TV advertising impressions demonstrate seasonality with significant annual ACF and partial ACF figures in all brigades. Magazine ad impressions reflect clear biannual peaks in ACF at 6- and 12-month lags. Radio advertising impressions are not consistent across the country. In Second and Third Brigades, the autocorrelation functions are significant at k=6. First Brigade has a significant positive ACF at a lag of 12 months. Fifth and Sixth Brigades have a significant positive ACF at both k=6 and k=12. Newspaper advertising impressions display no evidence of periodicity from the ACF.

In all regions except Sixth Brigade, Air Force high-quality male contracts demonstrate significant peaks in ACF at lags of 6 and 12 months. In the Sixth Brigade region, Air Force recruiting shows a significant peak in ACF only at 6 months, which is unexpected. The Marine Corps's production in this demographic reflects the strongest seasonal behavior of any service, with very significant ACF at lags of 12 and 24 months and, in all but the Third Brigade region, clearly significant ACF at a lag of 36 months. The ACFs' behavior also confirms the fact that the Marine Corps's high-quality male production appears the most consistent of all the services with little trend and very clear seasonality. In all but the Sixth Brigade region, the ACF for the Navy's high-quality male recruiting series demonstrates a much slower decay than the other services, indicating a less seasonal behavior. However, the Navy production correlograms does still have peaks in the ACF at a lag of k=12.

3. Correlation Between Time Series

The time series in each brigade are examined for high values of simple correlation between variables. Variables with high correlation present the potential for multi-collinearity, which can cause unstable models. The threshold for high correlation in this study is defined as $\rho \geq .95$. Identification of correlation values above this level do not represent a criterion for excluding variables from the initial model. However, the final reduced models are examined to ensure that both variables from pairs with high correlation values are not present.

In First Brigade, the correlation between high school graduate wages and the college attendance rate is 0.97. In Third Brigade, the correlation between these same variables is 0.95. These are the only two cases of correlation above the designated threshold. Both variables in this pair are from data that was transformed into monthly figures by linearizing between annual observations. In both First and Third Brigades, the college attendance rate variable is eliminated during the model reduction process.

4. Centered Data

In order to meet the requirement that an AR model must be developed for a zero mean series, the data for each variables is centered on the respective training set mean, $\bar{y}$ or $\bar{x}_i$. Not knowing in advance the order of the ARMA model that will be most effective, the centered data is used for all model development. The series averages and extrema for all variables in the full, test, and training data sets for each brigade are listed in Appendix A.

B. MODEL DEVELOPMENT

1. Initial Model Selection

Strictly autoregressive (AR), strictly moving average (MA), and mixed autoregressive moving average (ARMA) models are all explored during model development. Additionally, different manipulations of the centered data are explored including a one-period lead of all predictive variables and logistic transformations.

Moving average models using the centered data prove the most effective for creating models of the lowest order that have no significant ACF or PACF of the residuals.

The selection criteria for the moving average models are contained in Table 5.1. In this table, an asterisk in the "Moving Average Order" column indicates the model selected for step-wise reduction.

Moving Average Order (q)	Significant ACF or PACF of residuals	AIC
1st Brigade		
1 *	no	600.4
2	no	601.4
2nd Brigade		
1	no	639.4
2	no	634.0
3	no	633.5
4 *	no	623.8
5	no	628.1
3rd Brigade		
1	no	658.8
2 *	no	641.6
3	yes - PACF at lag $k = 8$	628.9
5th Brigade		
1	yes – ACF at lag $k = 2$, PACF at lag $k = 4$	589.5
2	yes - ACF at lag $k = 2$, PACF at lag $k = 4$	580.3
3	no	564.3
4	no	550.5
5 *	no	550.2
6	no	572.4
6th Brigade		
1	yes – PACF at lag $k = 7$	624.4
2	yes - PACF at lag $k = 7$	625.2
3	yes - PACF at lag $k = 4$	612.9
4	yes – ACF at lag $k = 4$,PACF at lag $k = 4$	599.5
5	yes - PACF at lag $k = 6$	605.5
6 *	no	591.3
7	no	592.8

Table 5.1 Model Selection Criteria Measures

The histograms contained in Figure 5.6 show that the assumption of a normal distribution of errors is plausible. First, Third, and Fifth Brigades' residual histograms are negatively skewed, while Second and Sixth Brigades' are positively skewed.

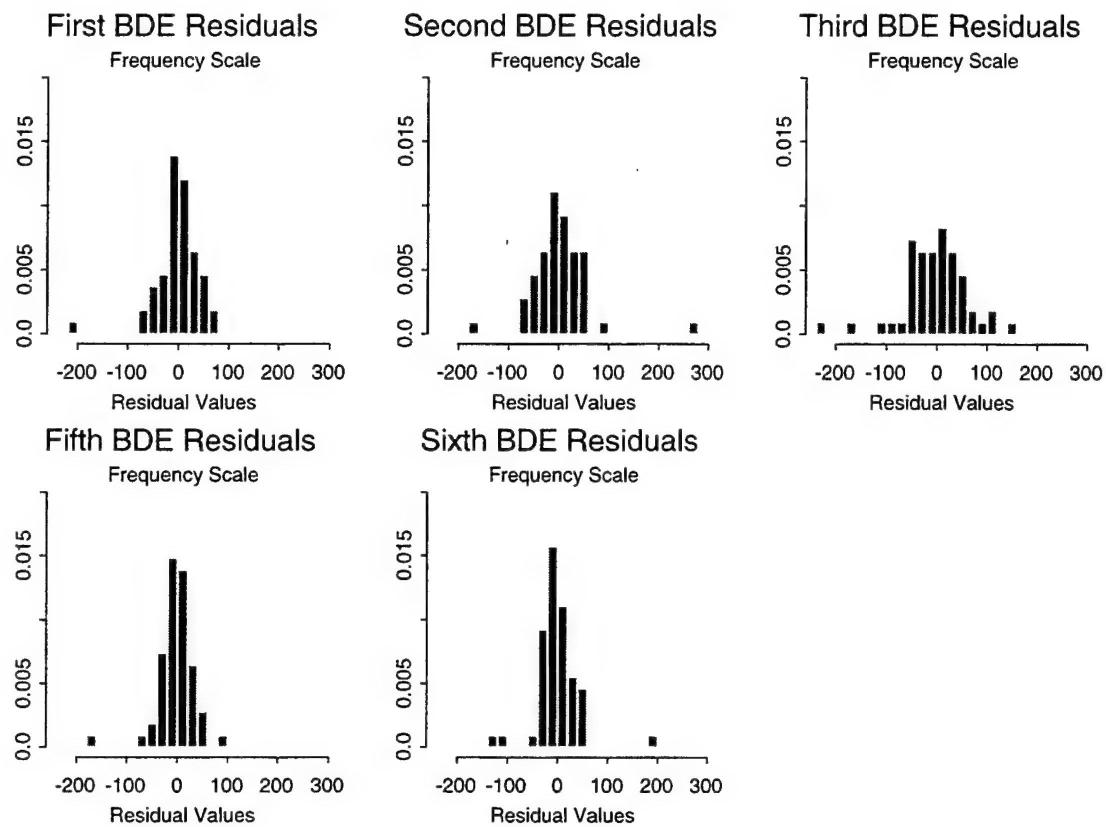


Figure 5.6 Residual Distributions for the Initial Models in Each Brigade

2. Stepwise Model Reduction

Reduced models are produced with the stepwise recursion using both the parametric and non-parametric means of sampling from the residual errors. The difference in which factors appear in the final models produced by each method is minimal. Only one variable in one of the five brigades differs between the two techniques. This result supports the supposition that the residuals from the initial models have normal distributions. The results of the non-parametric application of the stepwise reduction recursion are contained in Tables 5.3 and 5.4. The first table lists the sign of the

regressor and moving average coefficients of the initial full models and those of the significant factors in the reduced models. Table 5.4 lists the sign and magnitude of the regressor and moving average coefficients in the reduced models.

Final Models for Centered 54-month Data

Index	Factor	1st BDE Full Model	Reduced	2nd BDE Full Model	Reduced	3rd BDE Full Model	Reduced	5th BDE Full Model	Reduced	6th BDE Full Model	Reduced	# BDEs factor significant	# BDEs factor sig and +	# BDEs factor sig and -
1	MISSION	+	+	+		+	+	+	+	+		3	3	
2	RECRUITERS	-	-	+		+		+	+	-	-	3	1	2
3	TVADS													
4	RADIOADS													
5	MAGADS													
6	NEWSPADS													
7	COLOPTION	-		-	-	+		+	+	+	+	3	2	1
8	TGTPOP	+		+		-	-	-	-	-		2		2
9	UNEMP	+	+	-	-	-	-	-	-	-	-	5	1	4
10	HSGRADWAGE	+	+	-	-	-	-	-	-	+	+	5	2	3
11	GOTOCOLRATE	+		+	+	-		+		+	+	2	2	
12	COLLPREM	-	-	+		+	+	+		+	+	3	2	1
13	AFHQMC	-		+	+	-		-	-	+	+	3	2	1
14	MCHQMC	-	-	-		+	+	-	-	+	+	4	2	2
15	NHQMC	+	+	+		+	+	+	+	+	+	4	4	
16	FEB	-	-	-	-	+		-		-		2		2
17	MAR	-	-	-	-	+		-		-		2		2
18	APR	-	-	-	-	+		-	-	+		3		3
19	MAY	-	-	-	-	+		-	-	-		3		3
20	JUN	+	+	+	+	+	+	+	+	+	+	5	5	
21	JUL	+	+	+	+	-		+		-	-	3	2	1
22	AUG	+	+	+	+	-		+		-		2	2	
23	SEP	+	+	+	+	+		+	+	+	+	4	4	
24	OCT	+		+	+	-		-		+	+	2	2	
25	NOV	-		+		-		-	-	+		1		1
26	DEC	-	-	-	-	-	-	-	-	-	-	5		5
	MA 1	+	+	-	-	+	+	+	+	+	+			
	MA 2			+	-	+	+	+	+	+	+			
	MA 3			+	+			-	-	+	+			
	MA 4			+	+			+	+	+	+			
	MA 5							+	+	+	+			
	MA 6									-	-			
	# Significant factors per brigade		16		15		9		15		14			
	Non-month		7		5		7		9		9			
	Month		9		10		2		6		5			
	ARMA Model Order: (p,q)		(0, 1)		(0, 4)		(0, 2)		(0, 5)		(0, 6)			

Figure 5.3 Sign of Significant Factors and Moving Average Coefficients for Full and Reduced Models

Final Models for Centered 54-month Data

Index	Factor	1st BDE	2nd BDE	3rd BDE	5th BDE	6th BDE
1	MISSION	+ 1.76E-01		+ 1.96E+01	+ 2.77E-01	
2	RECRUITERS	- 3.23E-01			+ 4.59E-01	- 1.61E-01
3	TVADS					
4	RADIOADS					
5	MAGADS					
6	NEWSPADS					
7	COLOPTION		- 5.07		+ 5.77	+ 3.09
8	TGTPOP			- 0.00222	- 0.00987	
9	UNEMP	+ 6.28E+01	- 9.74E+01	- 2.65E+01	- 6.92E+01	- 4.35E+01
10	HSGRADWAGE	+ 9.10E+00	- 1.09E+01	- 2.94E+00	- 1.49E+00	+ 3.70E+00
11	GOTOCOLRATE		+ 675			+ 265
12	COLLPREM	- 5.08E+00		+ 1.15E+00		+ 1.90E+00
13	AFHQMC		+ 78.8		- 0.477	+ 0.532
14	MCHQMC	- 5.01E-01		+ 4.05E+01	- 4.52E-01	+ 5.05E-01
15	NHQMC	+ 3.13E-01		+ 4.02E+01	+ 1.22E+00	+ 5.55E-01
16	FEB	- 68.4	- 56.5			
17	MAR	- 4.95E+01	- 7.19E+01			
18	APR	- 1.37E+02	- 7.20E+01		- 9.06E+01	
19	MAY	- 1.46E+02	- 1.48E+02		- 5.13E+01	
20	JUN	+ 4.15E+01	+ 1.29E+02	+ 9.23E+01	+ 2.84E+02	+ 3.49E+01
21	JUL	+ 1.21E+02	+ 1.20E+02			- 2.99E+01
22	AUG	+ 1.63E+02	+ 6.09E+01			
23	SEP	+ 1.22E+02	+ 6.85E+01		+ 4.33E+01	+ 2.80E+01
24	OCT		+ 43.8			+ 30.3
25	NOV				- 62.6	
26	DEC	- 8.70E+01	- 1.18E+02	- 3.22E+01	- 9.19E+01	- 4.76E+01
	MA 1	+ 9.99E-01	- 3.92E-02	+ 3.94E-01	+ 1.55E+00	+ 2.24E-01
	MA 2		- 2.33E-06	+ 6.05E-01	+ 8.19E-01	+ 4.40E-01
	MA 3		+ 3.92E-02		- 1.96E+00	+ 6.23E-01
	MA 4		+ 9.99E-01		+ 1.80E-01	+ 4.40E-01
	MA 5				+ 4.15E-01	+ 2.24E-01
	MA 6					- 9.99E-01

Figure 5.4 Sign and Magnitude of the Reduced Model Regressors and Moving Average Coefficients

The total number of significant factors in each brigade varies between nine and sixteen. The number of significant continuous predictor variables in a brigade ranges from five to nine, while the number of significant monthly indicators varies between two and ten. All of the continuous variables are significant in at least two brigades, and all of the monthly indictors are significant in at least one brigade. Interpretation of the significant regressors is addressed in Section D.

3. Reduced Model Diagnostics

None of the residuals from the final reduced models display significant ACF or partial ACF. In all cases, the reduced models have a lower AIC value than the initial full models, as reflected in Table 5.5. Because the order of the models does not change during

the reduction process, the smaller AIC values stem from improvements in model accuracy.

	Initial Full Model AIC	Reduced Model AIC
1st Brigade	600	591
2nd Brigade	623	611
3rd Brigade	642	642
5th Brigade	550	546
6th Brigade	591	577

Table 5.5 Akaike Information Criteria Values for the Full and Reduced Models

C. FORECASTING

1. Nine-month Forecasts

The test set data and final reduced models are used in a simulation to produce a nine-month predicted time series of the centered response variable, z_t. The percent error of the forecast for each month, as defined in Chapter IV, is calculated and reflected in Table 5.6. The figures marked with an asterisk represent the cases in which the confidence interval contains the known response variable value.

BDE	Projected Month								
	1	2	3	4	5	6	7	8	9
1	-5.8*	-17.6	-22.4	-14.0	-16.8	-13.6	-1.8*	-8.2	-13.7
2	11.8*	4.8*	-14.7	-23.5	12.7*	23.4	11.4	8.1*	12.3
3	3.1*	-18.9*	-15.7*	-17.7*	-48.4	-22.7	-6.2*	-5.8*	-20.5
5	21.3*	13.0*	1.7*	-3.9*	-21.9*	-16.2*	19.5*	9.9	8.8*
6	-5.2*	18.2	18.1	15.7	1.0*	15.9	24.7	23.7	29.4

Table 5.6 Percent Error for Each Period of the Nine-Month Forecasts

The mean predicted value of the response variable, $\hat{y}_{t(forecast)}$, and a confidence interval of +/– 1 standard deviation of each forecast are plotted along with the known values of the response variable. The last three months of the training period and the predicted time series for each brigade are reflected in Figures 5.7 through 5.11. These graphs reveal that the forecasts do capture the general behavior of recruiting production during the test period.

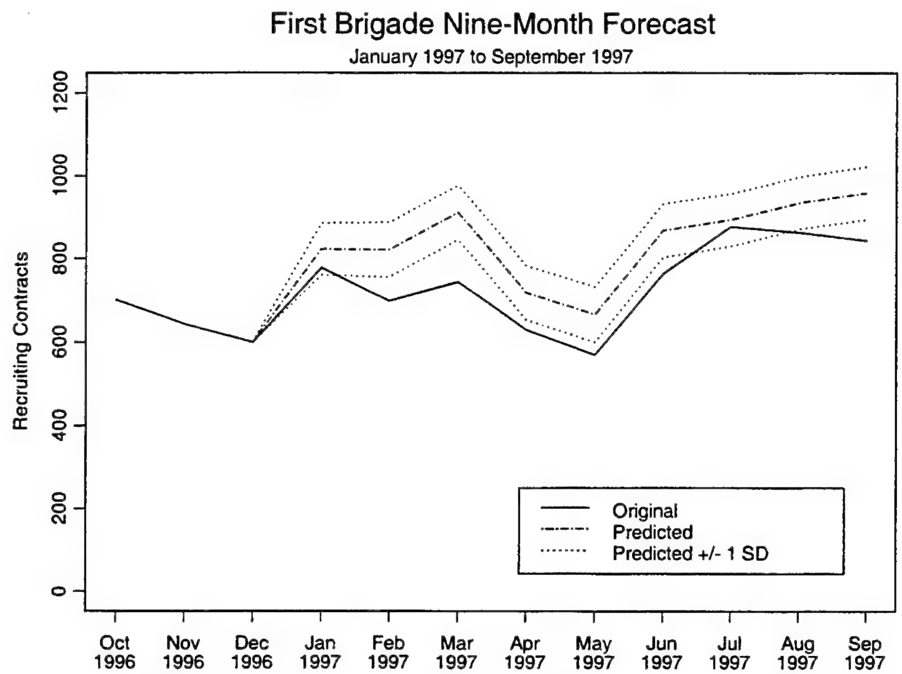


Figure 5.7 First Brigade Forecast

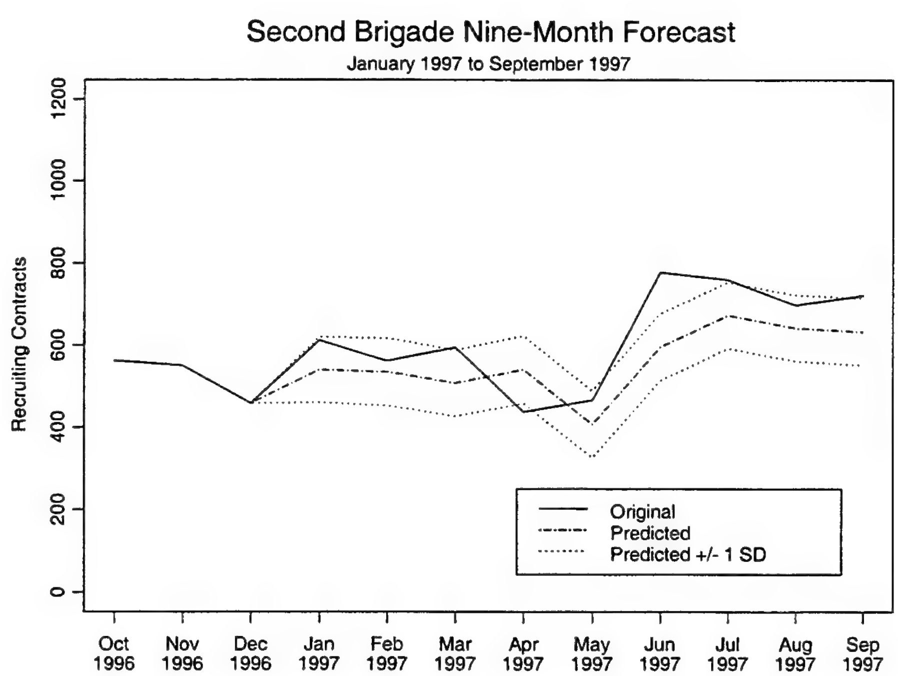


Figure 5.8 Second Brigade Forecast

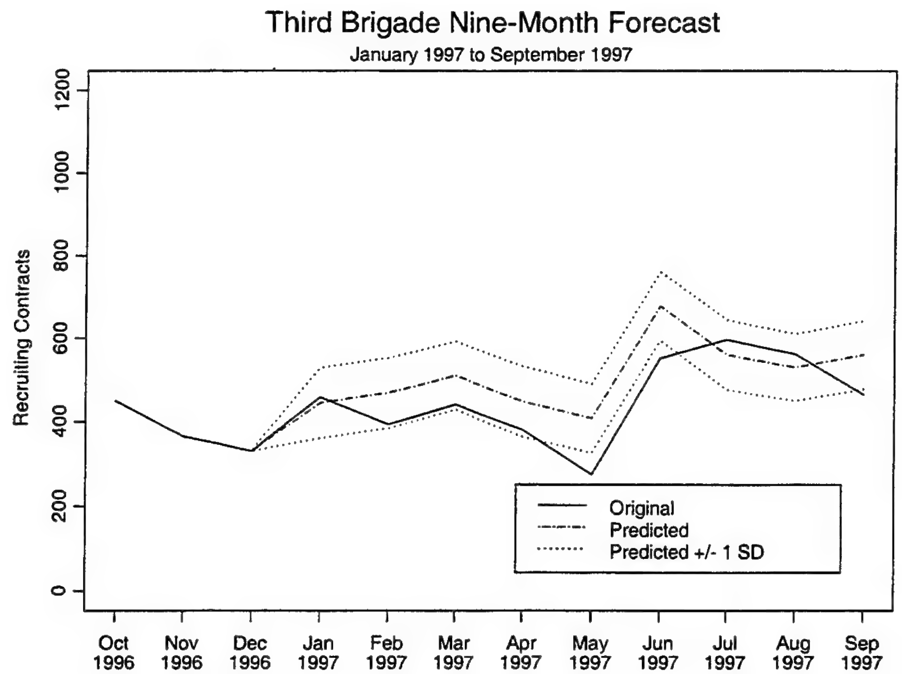


Figure 5.9 Third Brigade Forecast

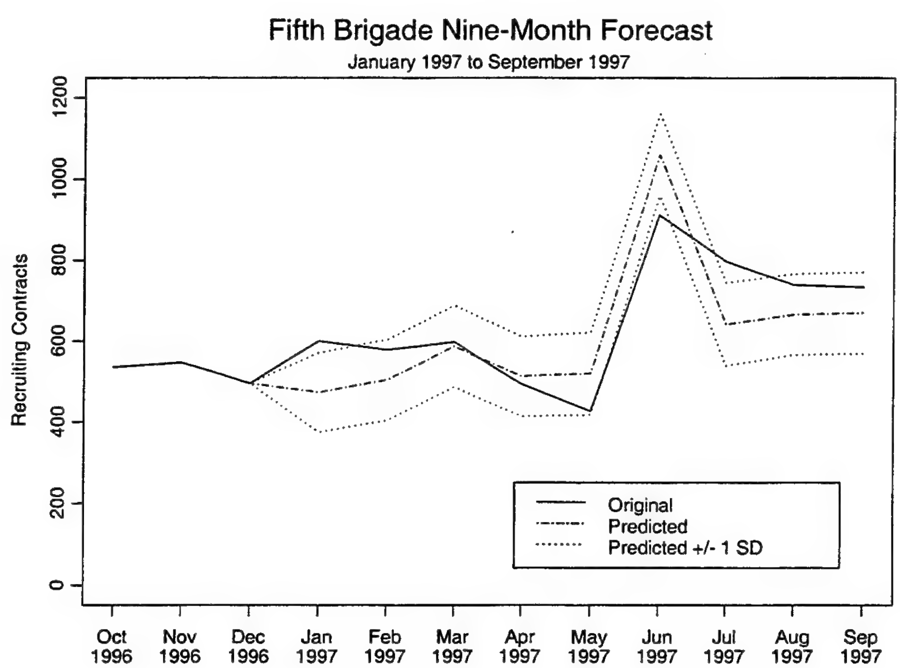


Figure 5.10 Fifth Brigade Forecast

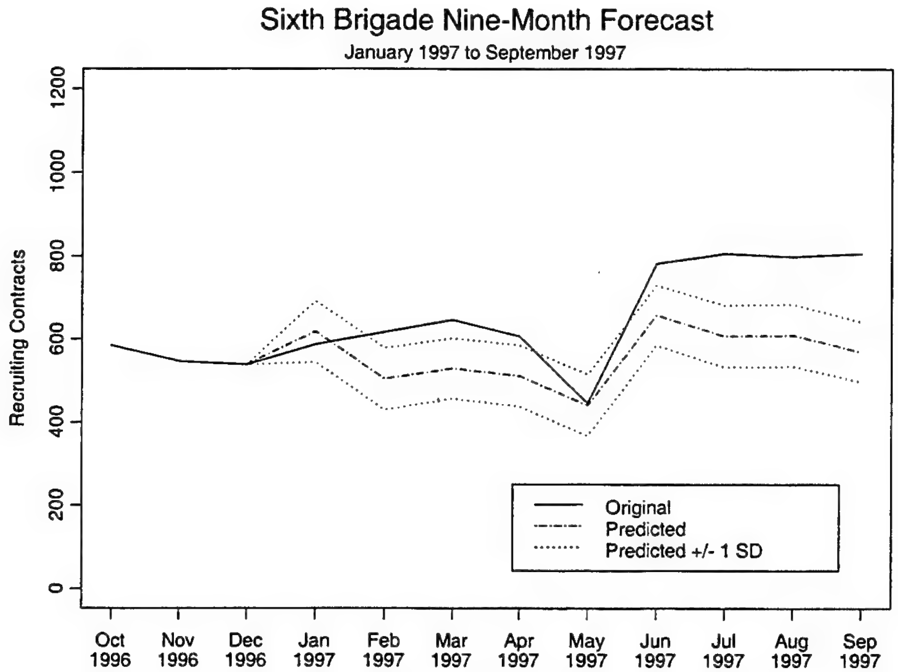


Figure 5.11 Sixth Brigade Forecast.

2. Forecast Error Diagnostics

The forecast errors, Δ_t, are calculated and plotted to determine if they appear randomly distributed. The forecast error plots are shown in Figure 5.12.

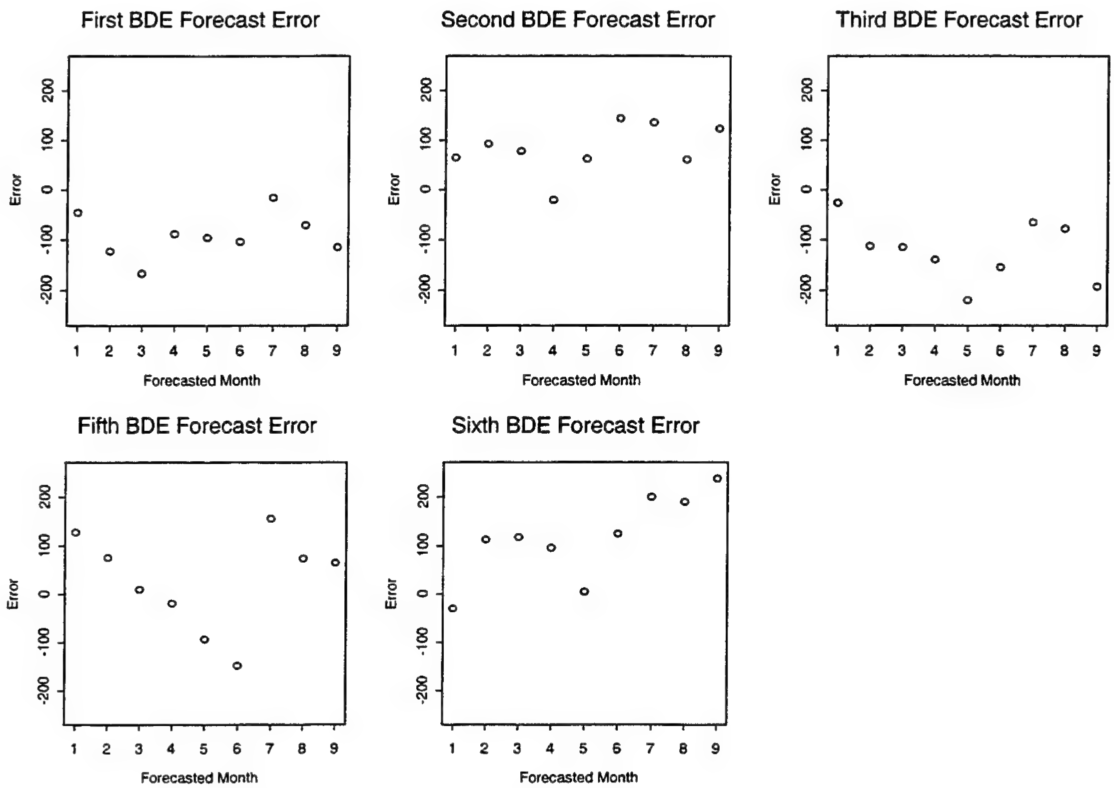


Figure 5.12 Plot of the forecast errors for each brigade

The limited number of forecasts makes it difficult to discern if the errors are randomly distributed. With the exception of Fifth Brigade, there are no apparent relationships that cause concern, such as errors increasing with forecast length or similarities in the distributions across brigades. The errors in Fifth Brigade demonstrate a near-linear decrease in the first six periods of the forecast series, but the last three periods appear to return to a random pattern. None of the forecast errors display significant ACF or partial ACF.

D. DISCUSSION

1. Regressor Coefficient Interpretation

Analysis of the significant factors in the final reduced models addresses three of the primary objectives of this thesis: validating factors from previous research; exploring recent suppositions regarding prominent factors; and enumeration of the differences in the recruiting environment throughout the country. In general, the presence of factors in the final models and the sign of their respective coefficients does not provide clear insight into the recruiting process. The significant factors vary considerably across brigades. In many cases, the signs of the coefficients for the same significant factors are inconsistent between brigades, and some are inconsistent with prior research. The significant factors and the regressor coefficient signs are listed in Tables 5.3.

At least one USAREC policy variable remains in the final model of each brigade. Mission is significant in three brigades and the coefficients' sign is positive as expected, meaning that recruiting production increases when USAREC issues higher quotas. Recruiter strength is significant in three brigades. In Fifth Brigade, increasing recruiter strength has a measurable positive effect on production. However, in First and Sixth Brigades, the sign of the coefficient for this variable is negative. This result initially appears counter-intuitive, but examination of the time series graphs reveals that, in these brigades, increases in recruiter strength did not provide the desired effect of boosting production. The number of eligible recruits receiving the college option is significant in three brigades. Like the behavior of the college option time series, the impact is not consistent, since the effect of the variable is positive in only two of these brigades.

The reduced models for each brigade contain at least two factors of the recruiting environment. The target population variable is significant in two brigades and for both the coefficient sign is negative. This result is counter-intuitive. Unemployment is a significant variable in all brigades, which is in line with prior research. However, the expected sign of the coefficient (positive) is present in only one brigade. Once again, this result may be partially explained by examining the behavior of the production and

unemployment time series. Over the period of this study, unemployment decreases in all regions, yet production remains fairly constant. The high school graduate wage level, which is also prominent in previous studies, is significant in all brigades. The variable coefficients display the expected sign (negative) in only three of the five brigades. The college attendance rate appears in the reduced models of two brigades, though the impact of the variable on recruiting production is positive and not negative as expected. The college premium variable is significant in three brigades and behaves as expected only in First Brigade.

At least one of the rival services' recruiting production variables appears in the final model of each brigade. Air Force high-quality male recruiting contracts are significant in three brigades, and are positive predictors in two of these regions. Marine Corps and Navy high-quality male recruiting variables are significant in four brigades. The Marine Corps production figures are negatively related to Army production figures in two brigades. The Navy production variable is the most consistent. Its behavior is positively related to Army production in all four brigades. In Sixth Brigade all rival service figures are significant and are positive predictors.

The sign of the coefficients for the monthly indicators behaves as expected in all but one case. December is consistently a difficult month, while June is a prolific month for recruiting. September is a significant and positive month in four of the brigades. In general, the winter months are negative, but not significant in all regions. All of these results are consistent with known recruiting production behavior. The negative coefficient for the July indicator in Sixth Brigade's final model represents the one counter-intuitive result for the binary monthly variables.

2. Model Form

It is difficult to draw conclusions regarding the first three thesis objectives because it is unclear whether the inconsistencies in the final models are due to random noise, inappropriate model form, new realities of the recruiting environment, and/or legitimate differences in the recruiting process across the country. The final models also produce disappointing results regarding the fourth thesis objective, which is to develop an

accurate tool for forecasting recruiting production. Chatfield provides some insight into a potential explanation. He states that "building a 'good' model from data subject to feedback can be difficult"(Chatfield, 1996). Feedback exists in systems in which the outputs in one period affect the inputs in future periods. In a discussion about econometric models, he states that in processes with rapid feedback loops, a single multivariate equation is less appropriate than modeling the system with multiple equations. However, he maintains that a multivariate model may still prove useful if the feedback in the system is slow and if the overall system is not well controlled by the inputs, as in the case of the economy.

Chatfield's observation may provide an explanation of why model accuracy improves when advertising data is removed. Clearly recruiting production results in one period have an impact on decisions about advertising expenditures in future periods. Concern about feedback raises the issue of whether all factors under USAREC's control should be eliminated from the model or whether multiple models need to be developed. Removal of all variables representing policies and resources under USAREC's control is unappealing because the resultant models would not provide any basis for analyzing policy and resource allocation. The feedback loop from production results to changes in recruiter strength is much slower than the response time in advertising feedback, which is an argument for maintaining this factor in model development. Based on the rapid feedback of production results on future missions and the ease with which missions can be changed, there may be a legitimate claim that inclusion of this variable has the potential to induce bias in the model coefficients. However, the reduced models that contain mission as a significant variable are no more or less accurate than those in which this factor is eliminated during the reduction process.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. CONCLUSIONS AND RECOMMENDATIONS

A. SUMMARY

The United States Army is currently experiencing the greatest recruiting shortfalls in the history of the All-Volunteer Force. The service faces unprecedented competition for young people as unemployment is at its lowest level in thirty years and college attendance rates are the highest in American history. Under these conditions, USAREC requires tools to quantify the impact of factors in the recruiting environment, to identify differences in the recruiting processes across its five regional subordinate units, and to measure the effectiveness of its policies and resource expenditures.

This thesis examined recruiting data from July 1992 to September 1997, a very dynamic period for the Army and Recruiting Command. The scope was limited to the high-quality male demographic, defined as a recruit who scored above the 50th percentile on the AFQT and who is a high school graduate or GED holder. This thesis aimed to do the following:

1. Validate factors from previous models on more current data.
2. Explore suppositions regarding new influences on the recruiting environment.
3. Enumerate the differences in the recruiting environment throughout the country.
4. Develop an accurate tool for predicting recruiting production.

Multivariate time series analysis was used to predict the number of enlistment contracts signed in a month as a function of exogenous and endogenous factors plus monthly indicators. Fifteen factors were initially included for examination in this study as predictive variables. They were selected based on their appearance in previous models or in recent research. The bootstrap was used to overcome the difficulties in determining significant factors presented by short duration of the recruiting data time series. This technique allowed resampling from within the existing data to provide robustness in the factor determination process. A stepwise recursion was developed to eliminate from the

time series models factors that were not statistically significant. The factors remaining in the reduced models were compared to those found to be significant in past research. The developed models were also used to create nine-month projections of recruiting production. The results were then compared to known production figures from test set data to determine forecast accuracy levels.

The final models indicated that unemployment figures and high school graduate wage levels are significant factors for predicting recruiting production. These results were consistent with findings from previous studies. However, the impact of these two factors was not clearly interpretable across the five recruiting brigades. In some brigades the effect of these variables on recruiting production was positive, while in other brigades the effect was negative. No consistent factors for measuring the competition between the Army and post-secondary schooling emerged from the model development process. The final models did successfully capture the seasonal nature of recruiting. There were considerable differences in the final model for each brigade, which despite probable noise in the data, indicated that influential predictors of recruiting production differ regionally. The forecasts produced using the final models captured the general behavior of the recruiting production series in the test period, but overall their accuracy was disappointing.

B. CONCLUSIONS

The 1990's were a dynamic period for U.S. Army recruiting. Predictions based on the past are dependent on the assumption that past patterns within each series and the relationships between series remain the same. Clearly, the recruiting process and environment was evolving rapidly over the period of this study. This evolution created noise in the data used in this thesis. Noise was likely also induced by the transformations required to convert some of the available data into a useable form for time series analysis. The results do support the intuition that the influential factors differ by region, a subject not addressed in the previous studies reviewed. Though the models developed in this

thesis may represent a descriptive tool for what occurred in the recruiting process during the period studied, they lack the forecasting accuracy to provide legitimate opportunities for "what if" analysis or optimization.

The most meaningful contribution of this thesis is the development of the stepwise recursion using bootstrap simulation for identifying significant factors in multivariate time series analysis. It proved to be a useful tool for providing robustness in a situation where data was limited. This methodology offers potential for further refinement and application.

C. RECOMMENDATIONS FOR FURTHER RESEARCH

Recommendations for improvements to the recruiting research fall in two categories: data collection and data handling. Future studies should attempt to have all predictive data specific to the targeted demographic. For example, college attendance rates, unemployment, and wage figures should address 17-24 year-old males only. All original data should consist of monthly observations. Additional indicator variables should be developed to represent additional policies, organizational structure, incentive programs, and specific events, such as military conflicts or government shutdowns. Multiple variables could be combined into single variable representations and co-integration vectors could be developed. Advanced time series analysis techniques should be explored including smoothing and filtering.

The stepwise reduction recursion using bootstrap simulation merits further exploration. A fist step should be testing the method with data for which accurate results have been determined through other techniques. An examination of the impact of changing the α values, the parameter for controlling the tolerance for which variables are removed from the model, is warranted. A method that uses incremental changes in the α values could be developed to rank order the significance of factors. Additionally, the number of model subsets examined in the model development process could be expanded by the creation of a forward addition stepwise recursion. These recommendations offer

means to further develop the potential shown by this technique for multivariate time series analysis.

APPENDIX A. ORIGINAL DATA SERIES SUMMARIES

First Brigade Data Summary

1ST Brigade

	AHQMC	MISSION	RECRUITERS	TVADS	RADIOADS	MAGADS	NEWSPADS	COLOPTION
Full Data Set								
Series Highs	1,020	2,428	1,491	4,021,949	8,849,750	3,557,303	221,846	34.60
Date	9209	9703	9709	9511	9709	9409	9703	9304
Series Lows	486	337	951	0	711,228	223,275	985	10.70
Date	9404	9511	9303	9307/9308/9409	9307	9307	9407	9702
Series Averages	704	1,111	1,116	1,881,163	4,733,137	2,068,248	44,071	18.46
Training Data Set								
Series Highs	1,020	1,586	1,249	4,021,949	3,557,303	3,557,303	221,846	34.60
Date	9209	9612	9509	9511	9409	9409	9703	9304
Series Lows	486	337	951	0	711,228	223,275	985	10.70
Date	9404	9511	9303	9307/9308/9409	9307	9307	9407	9402
Series Averages	695	963	1,078	1,755,268	4,546,288	2,102,080	41,589	19.19
Test Data Set								
Series Highs	877	2,428	1,491	3,542,808	8,849,750	2,498,510	221,846	20.00
Date	9707	9703	9707	9701	9709	9709	9703	9706
Series Lows	570	1,551	1,249	2,038,071	3,328,372	1,151,605	3,341	10.70
Date	9705	9701	9707	9708	9701	9707	9707	9702
Series Averages	752	1,997	1,340	2,636,531	5,854,233	1,865,255	58,965	14.06

	TGTPOP	UNEMP	HSGRADWAGE	GOTOCOLRATE	COLLPREM	AFHQMC	MCHQMC	NHQMC
Full Data Set								
Series Highs	528,050	8.22	479.00	37.00	281.00	606	628	899
Date	9207	9207	9709	9709	9709	9208	9207	9207
Series Lows	512,496	5.00	419.26	35.898	274.00	311	333	320
Date	9501	9612	9207	9207	9701	9412	9504	9705
Series Averages	516,870	6.22	447.12	36.44	250.89	392	456	508
Training Data Set								
Series Highs	528,050	8.22	469.00	36.90	273.40	606	628	899
Date	9207	9207	9612	9612	9612	9208	9207	9207
Series Lows	512,496	5.00	419.26	35.90	274.00	311	333	377
Date	9501	9612	9207	9207	9701	9412	9504	9511
Series Averages	516,636	6.36	442.54	36.35	246	394	454	518
Test Data Set								
Series Highs	519,495	6.03	479.00	37.00	281.00	438	572	568
Date	9709	9701	9709.00	9709	9709	9709	9707	9701
Series Lows	517,056	5.00	470.00	36.90	274.00	318	356	320
Date	9701	9709	9701	9701	9701	9705	9705	9705
Series Averages	518,275	5.43	474.60	36.95	277.82	381	468	448

Second Brigade Data Summary

	2ND Brigade							
	AHQMC	MISSION	RECRUITERS	TVADS	RADIOADS	MAGADS	NEWSPADS	COLOPTION
Full Data Set								
Series Highs	1,038	2,188	1,324	3,610,659	7,107,335	2,243,099	95,992	36.30
Date	9209	9706	9709	9511	9504	9409	9210	9304
Series Lows	410	443	854	0	477,477	129,618	146	11.20
Date	9405	9511	9211	9307/9308/9408	9307	9307	9510	9702
Series Averages	618	1,047	1,007	1,405,222	3,256,612	1,255,263	32,971	19.35
Training Data Set								
Series Highs	1,038	1,477	1,153	3,610,659	7,107,335	2,243,099	95,992	36.30
Date	9209	9612	9510	9511	9504	9409	9210	9304
Series Lows	410	443	854	0	477,477	129,618	146	11.20
Date	9405	9511	9211	9307/9308/9408	9307	9307	9510	9402
Series Averages	616	906	977	1,291,552	3,203,634	1,266,902	30,524	20.12
Test Data Set								
Series Highs	778	2,188	1,324	2,528,876	5,099,018	1,605,796	15,759	20.94
Date	9706	9706	9709	9701	9708	9709	9703	9706
Series Lows	438	1,420	1,420	1,450,087	2,468,390	842,530	8,384	11.20
Date	9704	9701	9701	9707	9702	9707	9707	9702
Series Averages	626	1,897	1,189	2,087,241	3,574,480	1,185,432	47,654	14.75

	TGTPOP	UNEMP	HSGRADWAGE	GOTOCOLRATE	COLLPREM	AFHQMC	MCHQMC	NHQMC
Full Data Set								
Series Highs	318,240	7.55	428.87	31.78	262.58	495	553	785
Date	9709	9207	9709	9708	9709	9508	9706	9207
Series Lows	288,101	4.40	367.42	31.28	222.16	254	286	276
Date	9301	9704	9306	9312	9406	9305	9305	9511
Series Averages	298,823	5.61	397.02	31.52	243.72	343	385	457
Training Data Set								
Series Highs	311,596	7.55	418.65	31.72	255.98	495	544	785
Date	9612	9207	9612	9612	9306	9508	9606	9207
Series Lows	288,101	4.94	367.42	31.28	222.16	254	286	276
Date	9301	9612	9306	9312	9406	9305	9305	9511
Series Averages	296,074	5.74	392.46	31.49	241.75	344	380	465
Test Data Set								
Series Highs	318,240	5.23	428.87	31.78	262.58	392	553	485
Date	9709	9701	9709	9708	9709	9709	9706	9706
Series Lows	312,392	4.40	419.78	31.73	248.50	292	326	287
Date	9701	9704	9701	9701	9701	9705	9704	9705
Series Averages	315,316	4.85	424.33	31.76	255.54	341	410	412

Third Brigade Data Summary

3RD Brigade

	AHQMC	MISSION	RECRUITERS	TVADS	RADIOADS	MAGADS	NEWSPADS	COLOPTION
Full Data Set								
Series Highs	924	1,561	1,025	3,300,763	6,800,640	2,935,059	652,669	34.66
Date	9209	9709	9709	9511	9504	9409	9202	9304
Series Lows	275	239	653	0	194,327	186,199	939	10.68
Date	9705	9511	9212	9307/9308/9408	9212	9307	9510	9702
Series Averages	471	814	810	1,508,527	3,319,025	1,673,852	57,322	18.48
Training Data Set								
Series Highs	924	1151	943	3,300,763	6,800,640	2,935,059	652,669	34.66
Date	9209	9609	9601	9511	9504	9409	9202	9304
Series Lows	290	239	653	0	194,327	186,199	939	10.70
Date	9511	9511	9212	9307/9308/9408	9212	9307	9510	9402
Series Averages	473	738	792	1,371,138	3,208,688	1,697,952	55,461	19.21
Test Data Set								
Series Highs	597	1,561	1,025	2,820,015	6,215,315	2,057,604	208,400	19.99
Date	9707	9709	9709	9701	9708	9709	9703	9706
Series Lows	275	979	852	1,668,449	2,064,227	1,035,668	10,141	10.68
Date	9705	9701	9701	9707	9701	9707	9707	9702
Series Averages	458	1,270	923	2,332,866	3,981,051	1,529,252	68,491	14.08

	TGTPOP	UNEMP	HSGRADWAGE	GOTOCOLRATE	COLLPREM	AFHQMC	MCHQMC	NHQMC
Full Data Set								
Series Highs	411,585	8.10	478.90	34.86	283.04	524	518	736
Date	9709	9201	9709	9709	9709	9208	9206	9206
Series Lows	392,008	3.72	400.71	33.78	194.66	226	230	233
Date	9401	9708	9201	9201	9206	9509	9504	9509
Series Averages	398,064	5.18	440.80	34.51	221.09	301	340	382
Training Data Set								
Series Highs	405,693	8.10	468.48	34.80	239.00	524	518	736
Date	9612	9201	9612.00	9612	9612	9208	9206	9206
Series Lows	392,008	3.98	400.71	33.78	194.66	226	230	233
Date	9401	9610	9201	9201	9206	9509	9504	9509
Series Averages	396,246	5.33	435.22	34.46	214.03	303	340	389
Test Data Set								
Series Highs	411,585	5.16	478.90	34.86	283.04	348	438	411
Date	9709	9701	9709	9709	9709	9709	9706	9707
Series Lows	406,368	3.72	478.91	34.86	283.04	228	231	270
Date	9701	9708	9709	9709	9709	9705	9705	9705
Series Averages	408,977	4.28	474.27	34.83	263.47	292	344	342

Fifth Brigade Data Summary

5TH Brigade

	AHQMC	MISSION	RECRUITERS	TVADS	RADIOADS	MAGADS	NEWSPADS	COLOPTION
Full Data Set								
Series Highs	1,000	1,789	1,236	3,727,596	8,007,866	2,806,682	227,944	34.56
Date	9209	9709	9709	9504	9503	9409	9703	9304
Series Lows	386	274	805	0	471,268	178,294	478	10.65
Date	9405	9511	9208	9307/9308/9408	9307	9307	9510	9702
Series Averages	624	1,011	959	1,699,711	4,048,646	1,644,373	45,791	18.43
Training Data Set								
Series Highs	1,000	1,455	1,101	3,727,596	8,007,866	2,806,682	227,944	34.56
Date	9209	9607	9510	9504	9503	9409	9703	9304
Series Lows	386	274	805	0	471,268	178,294	478	10.66
Date	9405	9511	9208	9307/9308/9408	9307	9307	9510	9402
Series Averages	620	907	938	1,518,629	3,870,762	1,662,282	43,535	19.16
Test Data Set								
Series Highs	912	1,789	1,236	3,648,237	6,986,570	2,119,179	227,944	19.93
Date	9706	9709	9709	9701	9709	9709	9703	9706
Series Lows	426	1,346	983	1,864,483	3,284,088	1,042,394	7,789	10.65
Date	9705	9701	9701	9707	9702	9702	9707	9702
Series Averages	653	1,637	1,087	2,786,203	5,115,948	1,536,913	59,327	14.04

	TGTPOP	UNEMP	HSGRADWAGE	GOTOCOLRATE	COLLPREM	AFHQMC	MCHQMC	NHQMC
Full Data Set								
Series Highs	317,489	7.23	419.13	33.50	241.01	509	544	784
Date	9709	9301	9709	9301	9709	9207	9306	9207
Series Lows	289,284	4.38	354.73	33.21	201.32	260	271	299
Date	9401	9705	9207	9709	9406	9511	9511	9511
Series Averages	298,177	5.54	383.20	33.43	219.21	348	383	511
Training Data Set								
Series Highs	309,803	7.23	409.29	33.50	237.87	509	544	784
Date	9612	9301	9612.00	9301	9306	9207	9306	9207
Series Lows	289,284	4.53	354.73	33.21	201.32	260	271	299
Date	9401	9610	9207	9709	9406	9511	9511	9511
Series Averages	295,524	5.66	377.94	33.46	216.52	349	381	514
Test Data Set								
Series Highs	317,489	5.61	419.13	33.32	241.01	389	521	632
Date	9709	9701	9709.00	9701	9709	9701	9706	9706
Series Lows	310,694	4.38	410.38	33.21	229.74	305	289	403
Date	9701	9705	9701	9709	9701	9705	9705	9705
Series Averages	314,092	4.85	414.76	33.26	235.38	344	397	487

Sixth Brigade Data Summary

6TH Brigade

	AHQMC	MISSION	RECRUITERS	TVADS	RADIOADS	MAGADS	NEWSPADS	COLOPTION
Full Data Set								
Series Highs	994	2,080	1,286	6,303,787	14,611,129	48,063,556	488,964	34.86
Date	9209	9709	9709	9504	9503	9409	9202	9304
Series Lows	434	345	878	0	604,769	302,848	0	10.66
Date	9511	9511	9406	9307/9308/9408	9307	9307	9510	9702
Series Averages	636	1,016	993	2,884,138	7,535,573	2,894,275	48,553	18.48
Training Data Set								
Series Highs	994	1419	1103	6,303,787	14,611,129	48,063,556	488,964	34.86
Date	9209	9607	9601	9504	9503	9409	9202	9304
Series Lows	434	345	878	0	604,769	302,848	0	10.68
Date	9511	9511	9406	9307/9308/9408	9307	9307	9510	9402
Series Averages	629	900	971	2,690,966	7,275,591	2,954,083	41,297	19.22
Test Data Set								
Series Highs	807	2,080	1,286	5,321,955	12,534,084	3,362,596	360,318	19.96
Date	9707/9709	9709	9709	9701	9708	9709	9703	9706
Series Lows	446	1,338	1,031	2,948,753	6,116,008	1,575,823	8,014	10.66
Date	9705	9701	9701	9707	9702	9707	9705	9702
Series Averages	678	1,711	1,125	4,043,167	9,095,465	2,535,424	92,085	14.06

	TGTPOP	UNEMP	HSGRADWAGE	GOTOCOLRATE	COLLPREM	AFHQMC	MCHQMC	NHQMC
Full Data Set								
Series Highs	419,163	8.98	448.99	40.01	283.17	479	544	781
Date	9709	9301	9709	9301	9709	9209	9206	9208
Series Lows	371,278	5.08	407.48	38.27	228.81	224	313	433
Date	9401	9709	9201	9709	9201	9310	9404	9705
Series Averages	386,302	6.81	428.96	39.08	257.54	351	417	564
Training Data Set								
Series Highs	406,135	8.98	438.78	40.01	282.30	479	544	781
Date	9612	9301	9612	9301	9612	9209	9206	9208
Series Lows	371,278	5.67	407.48	38.52	228.81	224	313	441
Date	9401	9612	9201	9612	9201	9310	9404	9405
Series Averages	381,794	7.01	426.38	39.19	253.33	351	416	565
Test Data Set								
Series Highs	419,163	6.47	448.99	38.49	283.17	413	520	634
Date	9709	9701	9709	9701	9709	9708	9707	9701
Series Lows	407,536	5.08	439.91	38.27	282.40	299	330	433
Date	9701	9709	9701	9709	9701	9702	9705	9705
Series Averages	413,349	5.59	444.45	38.38	282.79	346	422	561

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. S-PLUS CODE

Bootstrap Recursion S-Plus Code

```
function(mod, y, xreg, n = 250, parametric = F)
{
#
# bootstrapSim: Simulate ARIMA models for the boostrap
#
# args:     mod: arima model
#          xreg: matrix of regression dependent variables
#             n: Number of trials
#   parametric: Use parametric (Normal-based) bootstrap? Default F.
#
#
      if(class(mod) != "arima") stop("This function operates on an
       arima model")
#
# Set up output
#
      this.is.ar <- this.is.ma <- F
      xreg.out <- matrix(NA, n, length(mod$reg.coef))
      if(any(names(mod$model) == "ar")) {
            this.is.ar <- T
            ar.out <- matrix(0, n, mod$model$order[1])
      }
      if(any(names(mod$model) == "ma")) {
            this.is.ma <- T
            ma.out <- matrix(0, n, mod$model$order[3])
      }
#
# Extract residuals
#
      resids <- arima.diag(mod, resid = T, plot = F)$resid
      r.len <- length(resids) #
#
# If data is missing, try to find it.
#
      if(missing(xreg)) {
            name <- as.character(mod$call)[4]
            if(!exists(name))
                  stop(paste("Can't find regressors in", name))
            xreg <- get(name)
      }
      if(missing(y)) {
            name <- mod$series
            if(!exists(name))
                  stop(paste("Can't find y data in", name))
            y <- get(name)
      }
#
#
```

```
# Main loop
#
      if(mod$model$order[1] != 0) {
            skip <- mod$model$order[1]
            first.resids <- y[1:skip] - xreg[1:skip,  ] %*%
              mod$reg.coef
#
# the previous step accounts for the start-up cost of AR models and
# fills in the p missing residuals with approximations.
#
            for(i in 1:n) {
                  if(i %% 100 == 0) {
                        cat("Operating on loop ", i, "\n")
                  }
                  if(parametric) {
                        cat("parametric\n")
                        resid.sd <- sqrt(var(resids))
                        new.y <- arima.sim(mod$model, xreg = xreg,
                          reg.coef = mod$reg.coef, innov
                          = rnorm(n = length(resids), sd = resid.sd))
                  }
                  else {
                        new.resids <- c(first.resids,
                          resids[sample((skip + 1):r.len, replace =
                          T)])
                  }
                  new.y <- arima.sim(mod$model, xreg = xreg, reg.coef =
                    mod$reg.coef, innov = new.resids)
                  new.model <- arima.mle(new.y, model = mod$model, xreg
                    = xreg, max.fcal = 400, max.iter = 250)
                  if(new.model$converged == F) {
                        cat("Warning: model", i, "didn't converge!\n")
                  }
                  else {
                        if(this.is.ar)
                          ar.out[i,  ] <- new.model$model$ar
                        if(this.is.ma)
                          ma.out[i,  ] <- new.model$model$ma
                        xreg.out[i,  ] <- new.model$reg.coef
                  }
            }
      }
      else {
            first.resids <- y - xreg %*% mod$reg.coef
            for(i in 1:n) {
                  if(i %% 100 == 0) {
                        cat("Operating on loop ", i, "\n")
                  }
                  new.resids <- resids[sample(1:r.len, replace = T)]
                  if(parametric) {
                        cat("parametric\n")
                        resid.sd <- sqrt(var(resids))
```

```
                        new.y <- arima.sim(mod$model, xreg = xreg,
                          reg.coef = mod$reg.coef, innov
                          = rnorm(n = length(resids), sd = resid.sd))
#
# If "innov" is supplied, it should be a vector,
# e.g. innov = rnorm (n = length(resids), sd = resid.sd)
# If "innov" is NOT supplied, then a vevtor of innovations is generated
# by the rand.gen() function, which, by default, is rnorm. Additional
# arguments to this function can be pased as arguments to arima.sim,
# e.g. [innov = not passed], rand.gen = rnorm, n = length(resids),
# sd = resid.sd
#
                  }
                  else {
                        new.y <- arima.sim(mod$model, xreg = xreg,
                          reg.coef = mod$reg.coef, innov = new.resids)
                  }
                  new.model <- arima.mle(new.y, model = mod$model,
                    xreg = xreg, max.fcal = 300, max.iter = 150)
                  if(new.model$converged == F) {
                        cat("Warning: model", i, "didn't converge!\n")
                  }
                  else {
                        if(this.is.ma)
                          ma.out[i,  ] <- new.model$model$ma
                        xreg.out[i,  ] <- new.model$reg.coef
                  }
            }
      }
      if(this.is.ar)
            if(this.is.ma)
                  return(list(Xreg = xreg.out, AR = ar.out,
                    MA = ma.out))
            else return(list(Xreg = xreg.out, AR = ar.out))
      else return(list(Xreg = xreg.out, MA = ma.out))
}
```

Stepwise Reduction Recursion S-Plus Code

```
function(mod, y, regressors, n = 1000, parametric = F, SD.range = 1,
  maximumIterations = 1)
{
#
# Stepwise: eliminate time-series regressors by backward elimination.
#
# Arguments:     mod: arima model
#                  y: y data vector(response variable)
#         regressors: matrix of regressors
#                  n: n to be passed to bootstrapSim
#         parametric: to be passed to bootstrapSim
#           SD.range: a tolerance for deciding when to stop deleting
#                     columns. Stop deleting regressors when no column
#                     of regression coefficients has "0" in the range
#                     (mean +/- SD.range * SD).
# maximumIterations: maximum number of times to run through the
#                     discarding loop. Default is 1.
#
#
      stillDiscarding <- T
      counter <- 0       #
#
# Save "y" in frame 1 so "arima.diag" and others can find it if needed
#
      assign("y", y, frame = 1)
      while(stillDiscarding && counter < maximumIterations) {
            counter <- (counter + 1)
#
# Print the current columns and run bootstrapSim.
#
            cat("Loop ", counter, ", cols are ",
              dimnames(regressors)[[2]], "\n")
            cat("Calling bootstrapSim to execute bootstrap \n")
            bsRegCoefs <- bootstrapSim (mod, y, regressors, n = n,
              parametric = parametric)$Xreg
            proportion <- vector("single", ncol(bsRegCoefs))
#
# For each column of the resulting bootstrap regression coefficients,
# find the proportion of values that fall between 0 and (2 * the mean
# of the column). The objective is to discard the column with the
# largest such proportion; that's the column in which most of the
# values are close to 0. Each column of bsRegCoefs corresponds to a
# factor in the time series.
#
            for(i in 1:(ncol(bsRegCoefs) - 1)) {
                  cat("Examining factor ",
                    dimnames(regressors)[[2]][i], "\n")
                  numerator <- 0
                  x.bar <- mean(bsRegCoefs[, i], na.rm = T)
                  if(na.sum <- sum(is.na(bsRegCoefs[, i])))
```

```
                        cat(paste("Encountered", na.sum, "missing
                          coeffs\n"))
                sd <- sqrt(var(bsRegCoefs[, i], na.method = "omit"))
#
# If 0 is in the range of the mean +/ SD.range * sd, calculate
# proportion. If not, assign that column a proportion of 0.
#
                if(x.bar - SD.range * sd <= 0 && x.bar + SD.range *
                  sd >= 0) {
                        for(j in 1:n) {
                                if(bsRegCoef[j, i] == !NA) {
                                        if(bsRegCoefs[j, i] < SD.range * sd
                                          && bsRegCoefs[j, i] > - (SD.range
                                          * sd)) {
                                                numerator <- numerator + 1
                                        }
                                }
                        }
                        proportion[i] <- numerator/n
                        cat("Proportion is ", proportion[i], "\n")
                }
                else proportion[i] <- 0
        }
#
# After examining all factors, choose the factor which has the highest
# proportion. If all proportions are equal to 0, set flag to exit
# "while" loop.
#
        if(all(proportion == 0)) {
                cat("No additional factors discarded \n")
                stillDiscarding <- F
        }
        else {
                max.index <- (1:length(proportion))[proportion ==
                  max(proportion)][1]
                cat("Discarding ",
                  dimnames(regressors)[[2]][max.index], "\n")
                regressors <- regressors[,  - max.index]
                assign("regressors", regressors, frame = 1)
                mod <- arima.mle(y, model = list(order =
                  mod$model$order), xreg = regressors, maxfcal = 300,
                  max.iter = 150)
        }
    }
    cat("final significant factors are: ", dimnames(regressors)[[2]],
      "\n")
    return(regressors)
}
```

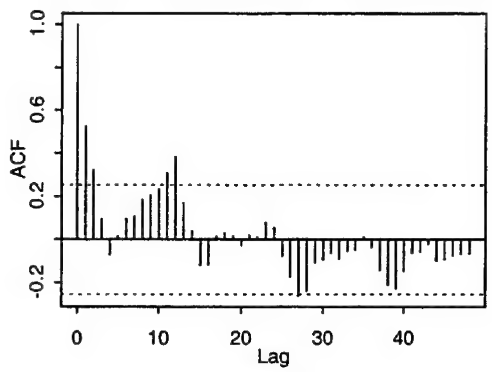
Forecasting Simulation S-Plus Code

```
function(mod, N, XREG, regressionCoef, loops, ...)
{
#
# makePrediction: generates multiple ARIMA simulations to predict a
#                 univariate time series. Returns the mean and sd of
#                 the predictions for each period in the series.
#                 Also returns a histogram of the simulated values for
#                 the first period in the predicted time series. The
#                 ... notation allows the user to specify how the
#                 innovations for arima.sim are created.
#
# args:        mod: the order for the ARIMA model
#                N: the number of periods in the desired time series
#             XREG: a matrix of regression variable values
# regressionCoef: a vector of regression coefficients corresponding to
#                 xreg
#
#
      predOut <- matrix(nrow = loops, ncol = 9)
      for(i in 1:loops) {
            x <- arima.sim(model = mod, n = N, xreg = XREG,
                           reg.coef = regressionCoef, ...)
            predOut[i,  ] <- x
      }
      mean <- apply(predOut, 2, mean)
      var  <- apply(predOut, 2, var)
      sd   <- sqrt(var)
      hist(predOut[, 1], nclass = 20)
      return(mean, sd)
}
```

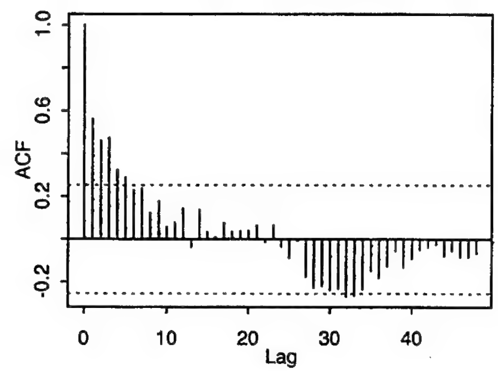
APPENDIX C. ORIGINAL DATA SERIES CORRELOGRAMS

First Brigade Autocorrelation Function Correlograms

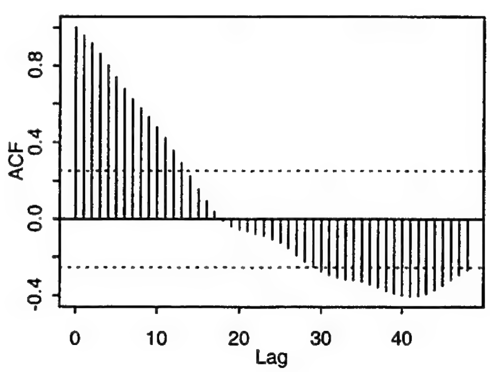
H-Q Male Recruiting Production



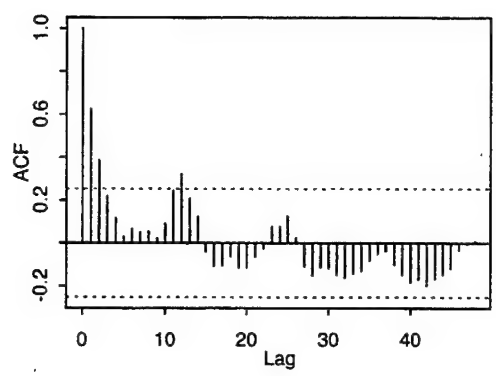
H-Q Male Recruiting Mission



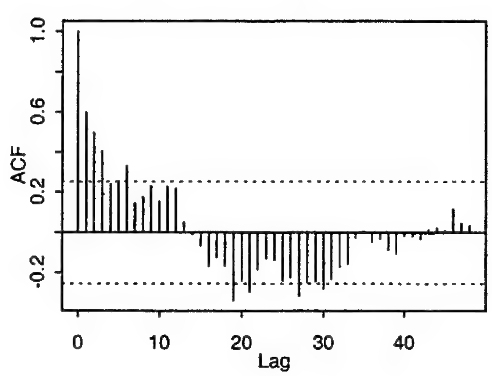
Recruiter Strength



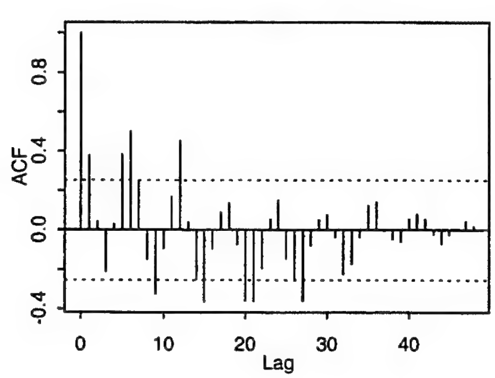
TV Ad Impressions



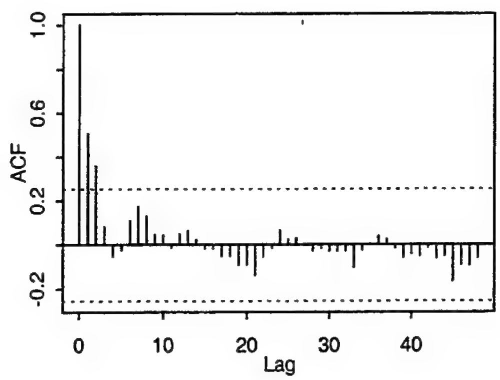
Radio Ad Impressions



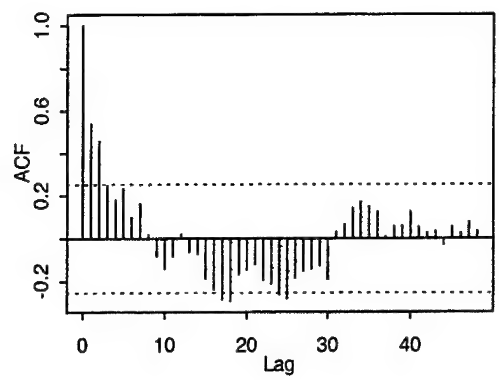
Magazine Ad Impressions



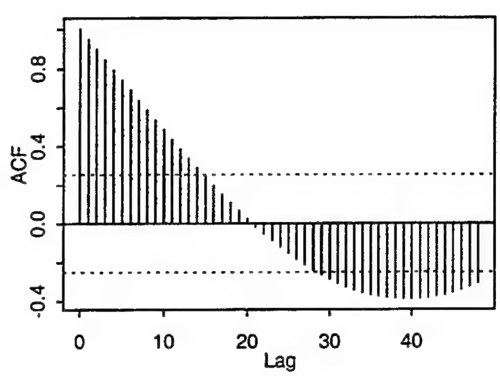
Newspaper Ad Impressions



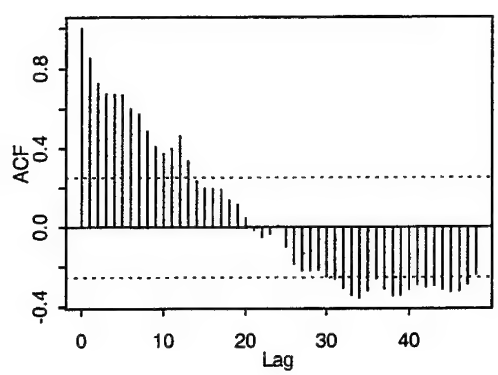
% Receiving College Fund



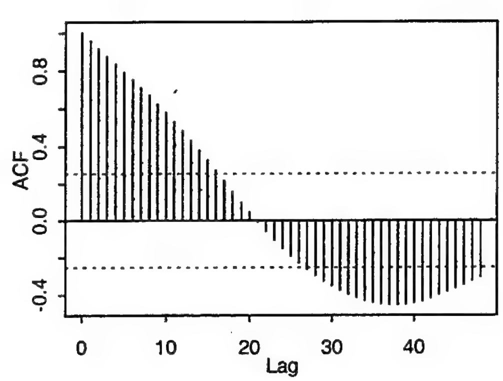
Target Population



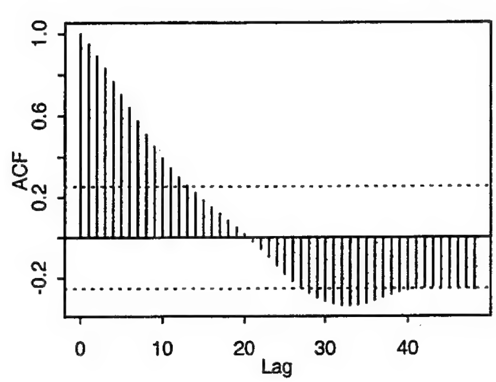
Unemployment Rate

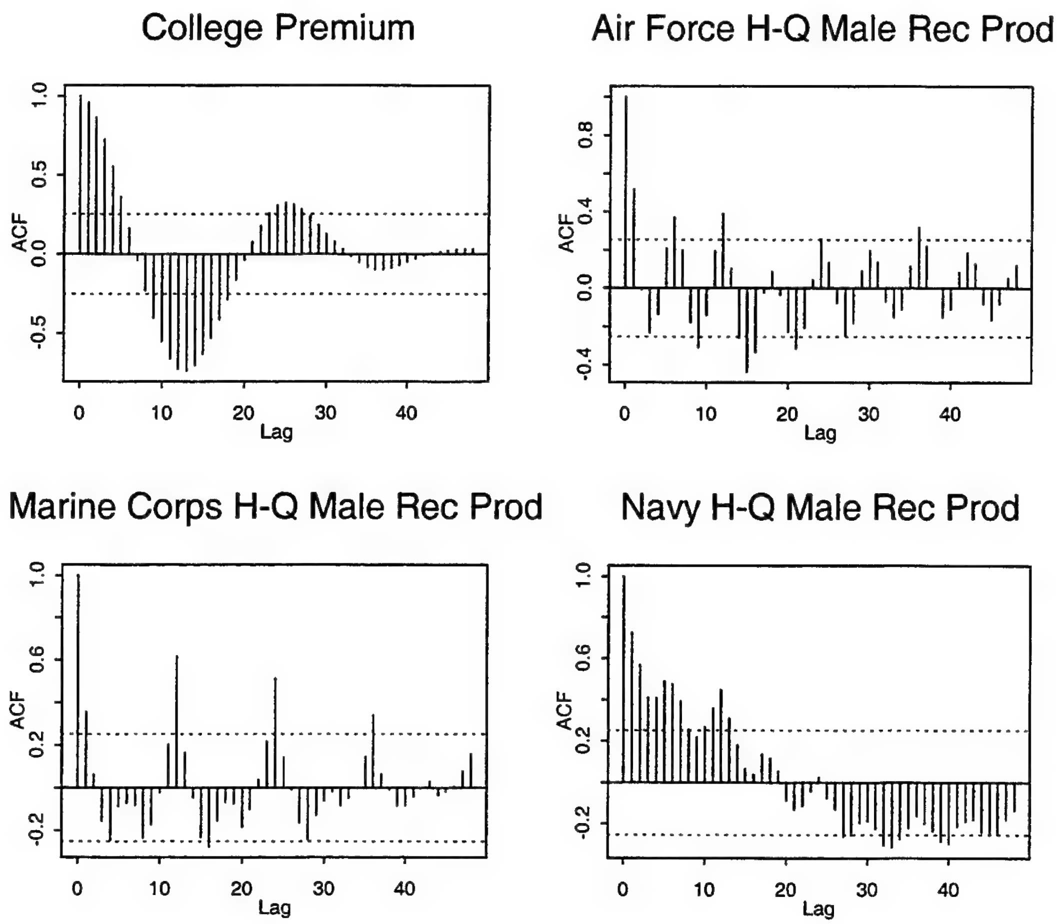


High School Grad Wage Level



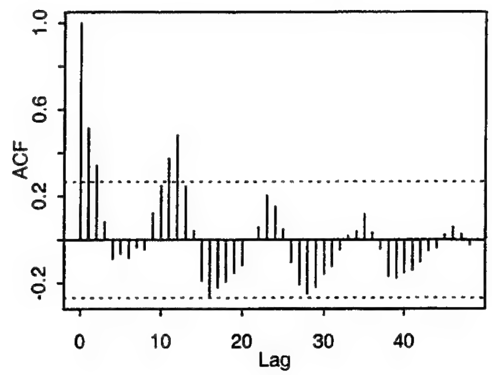
College Attendance Rate



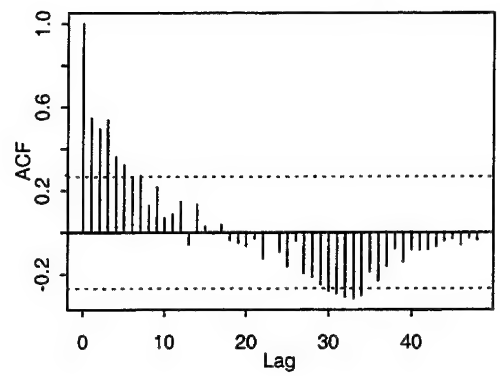

College Premium
Air Force H-Q Male Rec Prod
Marine Corps H-Q Male Rec Prod
Navy H-Q Male Rec Prod
ACF
Lag

Second Brigade Autocorrelation Function Correlograms

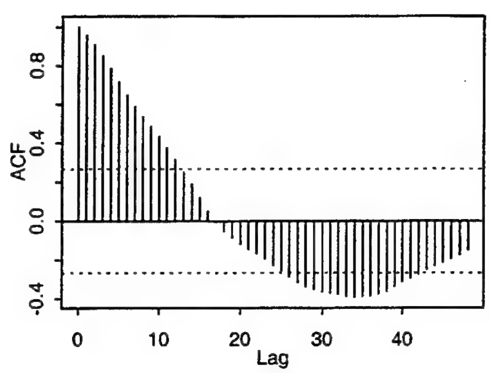
H-Q Male Recruiting Production



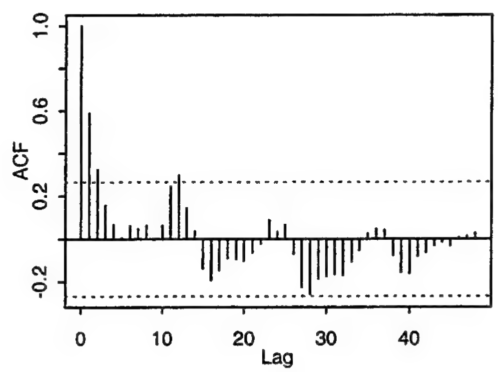
H-Q Male Recruiting Mission



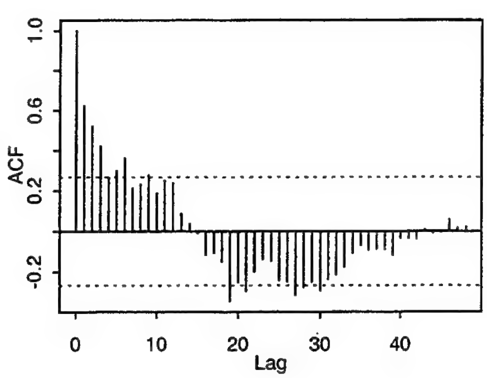
Recruiter Strength



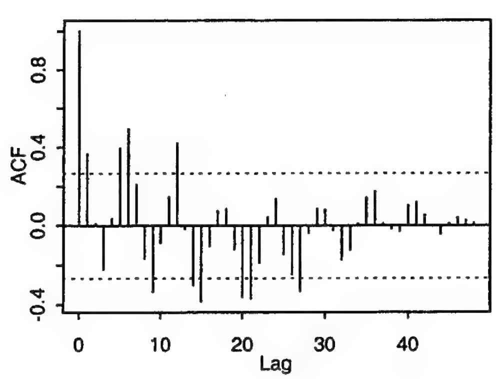
TV Ad Impressions

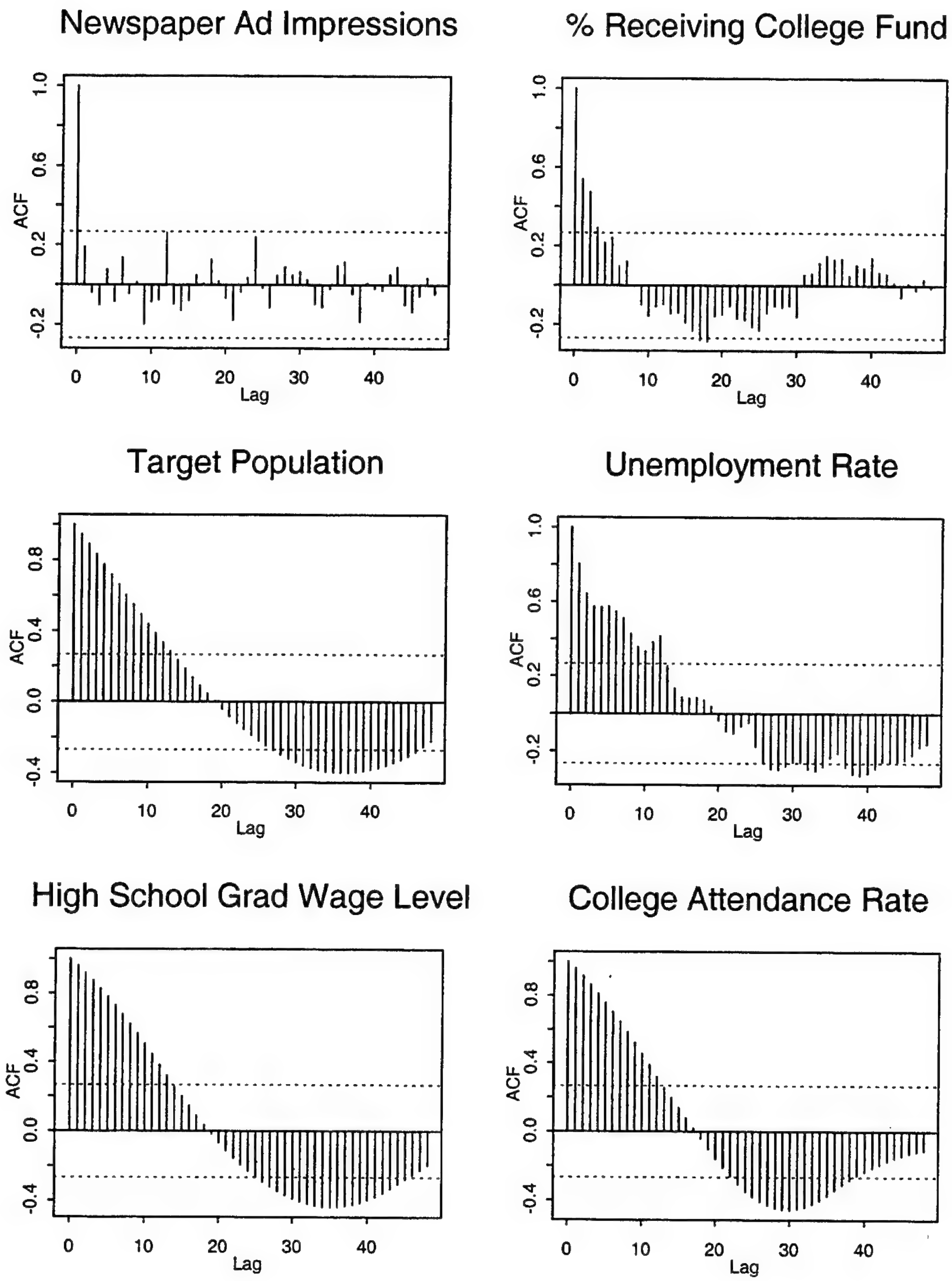


Radio Ad Impressions



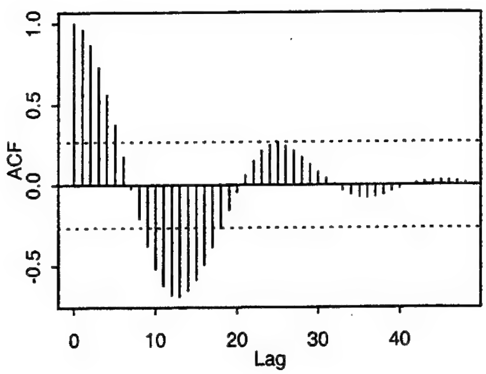
Magazine Ad Impressions



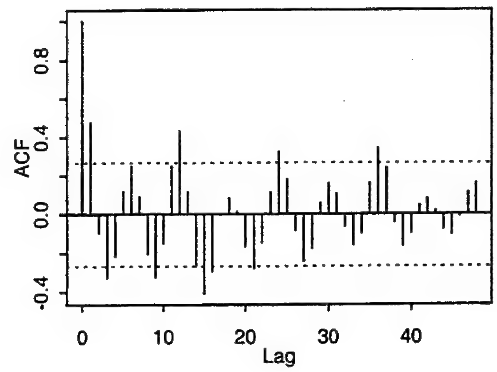


Newspaper Ad Impressions
% Receiving College Fund
Target Population
Unemployment Rate
High School Grad Wage Level
College Attendance Rate
ACF
Lag

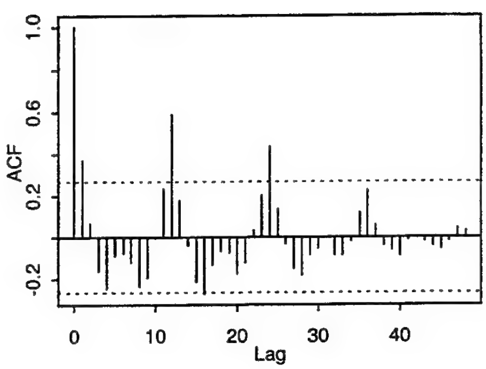
College Premium



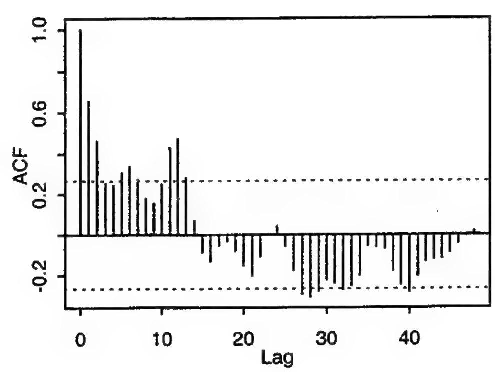
Air Force H-Q Male Rec Prod



Marine Corps H-Q Male Rec Prod

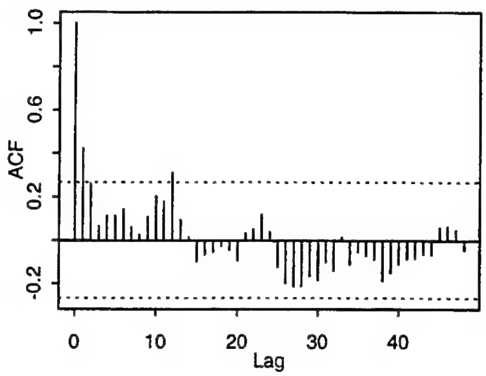


Navy H-Q Male Rec Prod

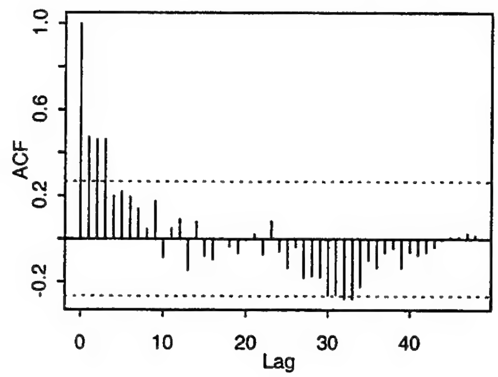


Third Brigade Autocorrelation Function Correlograms

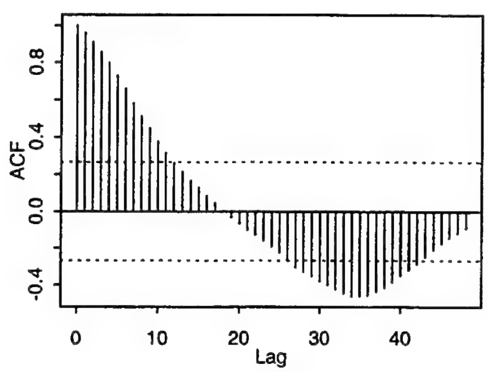
H-Q Male Recruiting Production



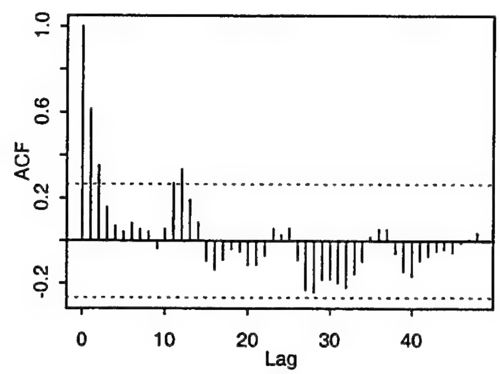
H-Q Male Recruiting Mission



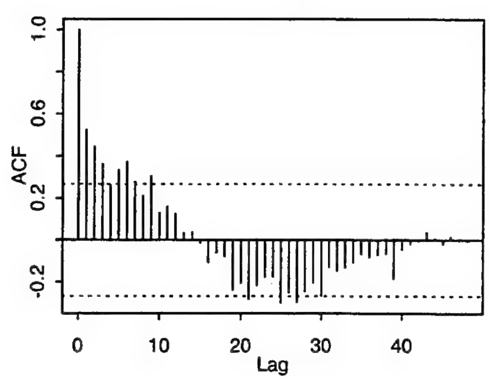
Recruiter Strength



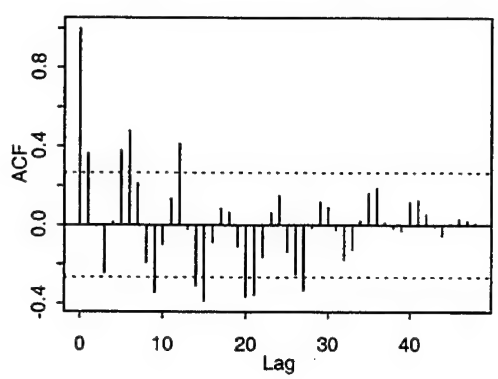
TV Ad Impressions



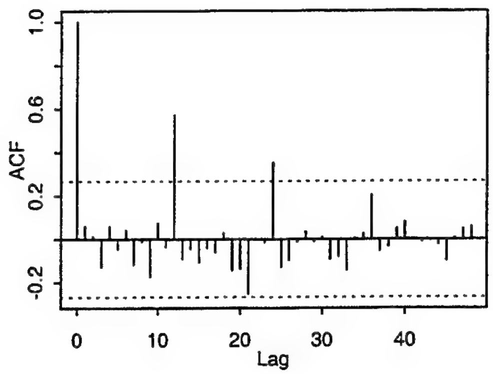
Radio Ad Impressions



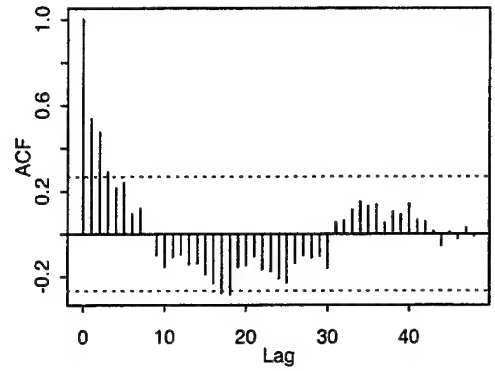
Magazine Ad Impressions



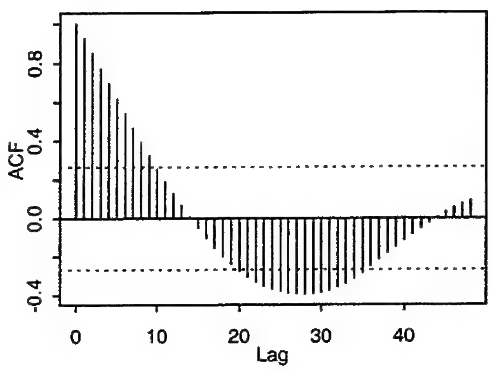
Newspaper Ad Impressions



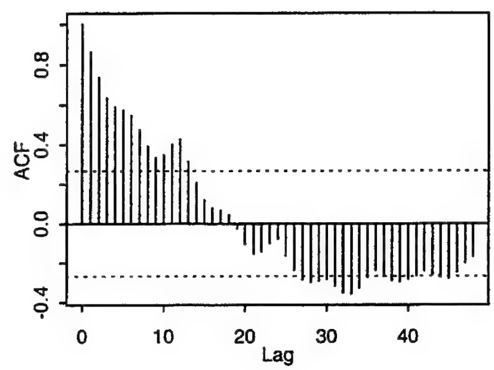
% Receiving College Fund



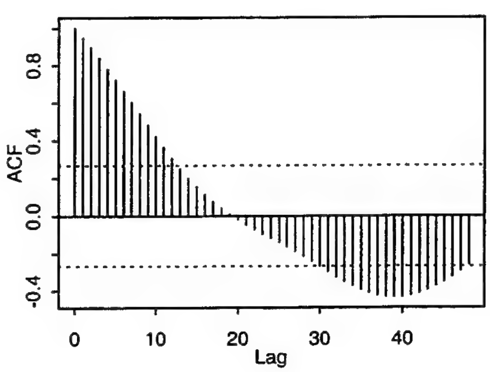
Target Population



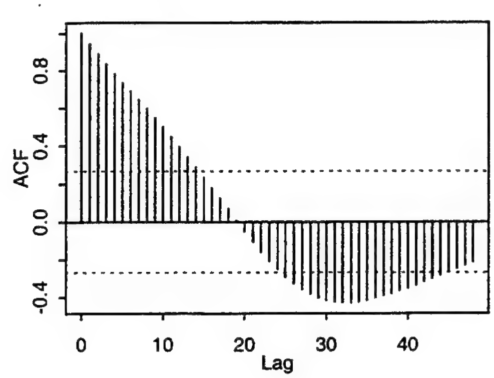
Unemployment Rate

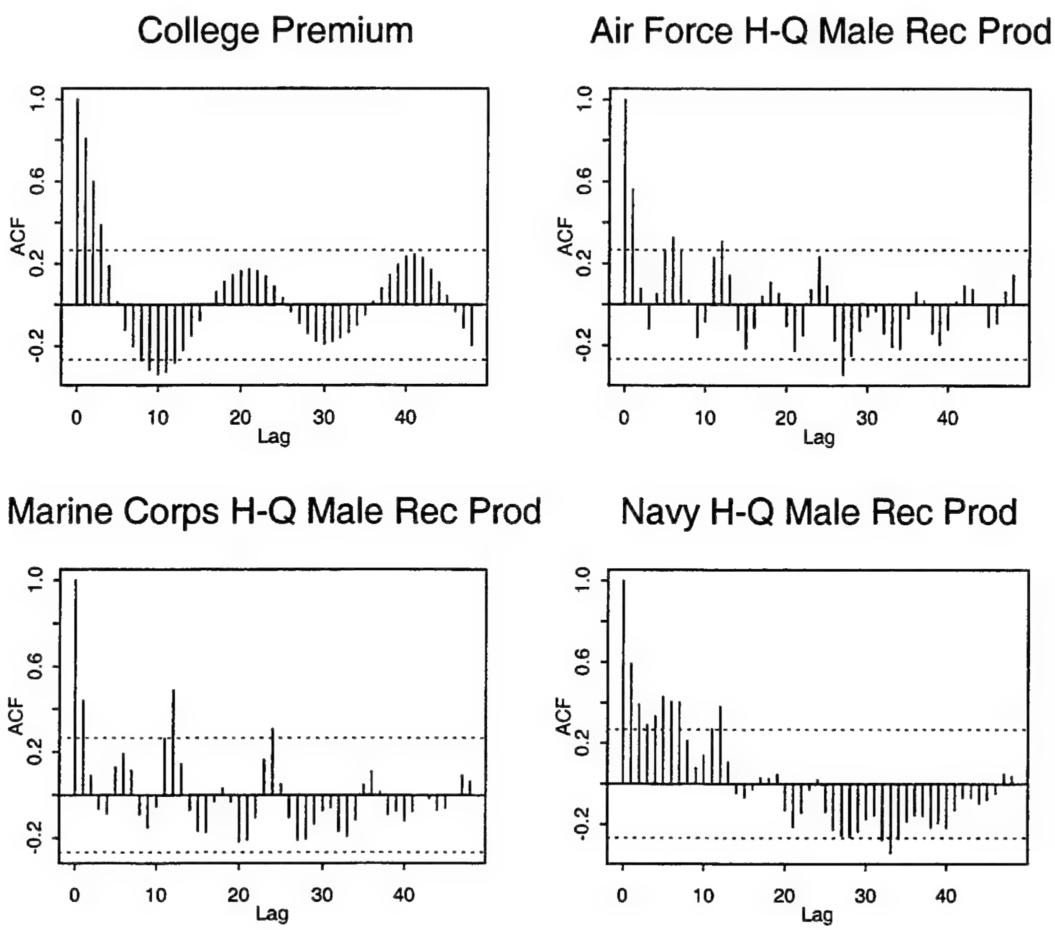


High School Grad Wage Level



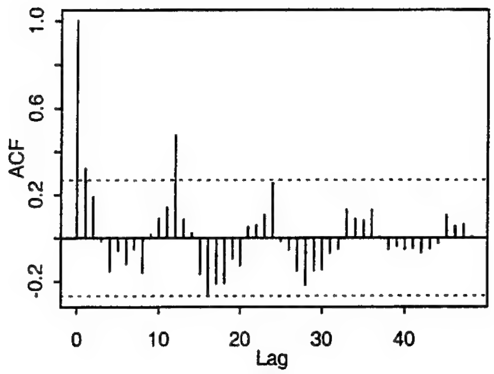
College Attendance Rate



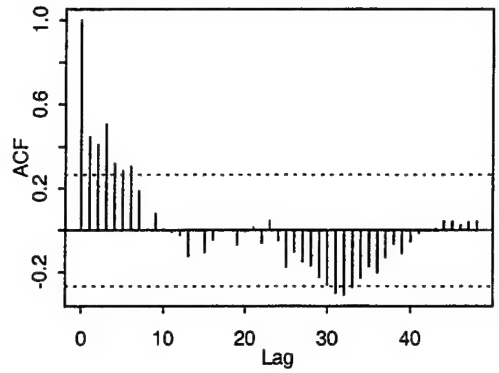

College Premium
Air Force H-Q Male Rec Prod
Marine Corps H-Q Male Rec Prod
Navy H-Q Male Rec Prod
ACF
1.0
0.6
0.2
-0.2
0
10
20
30
40
Lag

Fifth Brigade Autocorrelation Function Correlograms

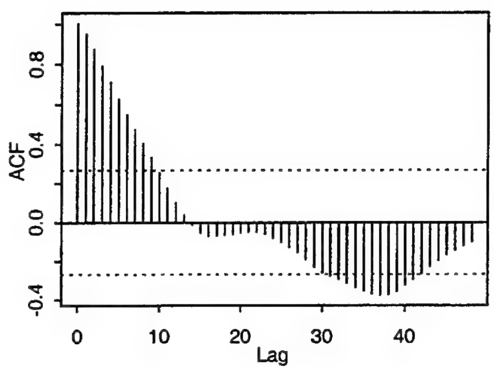
H-Q Male Recruiting Production



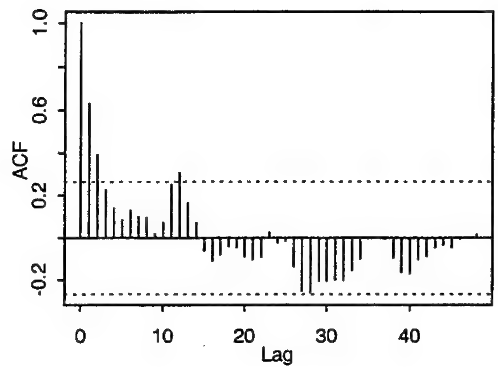
H-Q Male Recruiting Mission



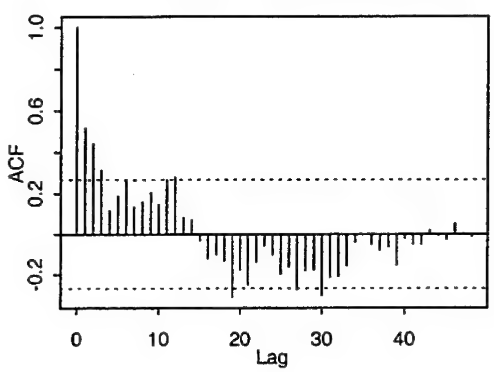
Recruiter Strength



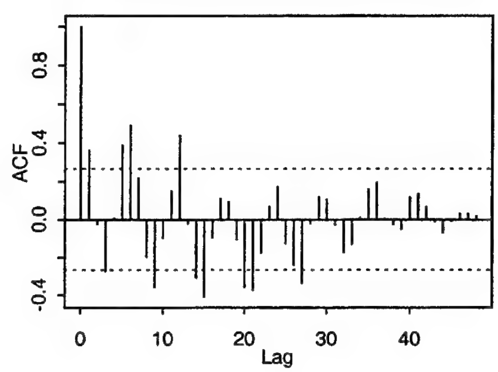
TV Ad Impressions



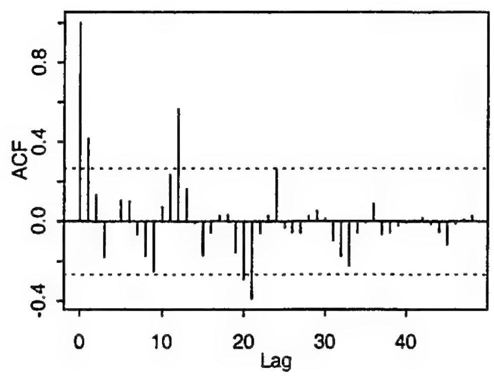
Radio Ad Impressions



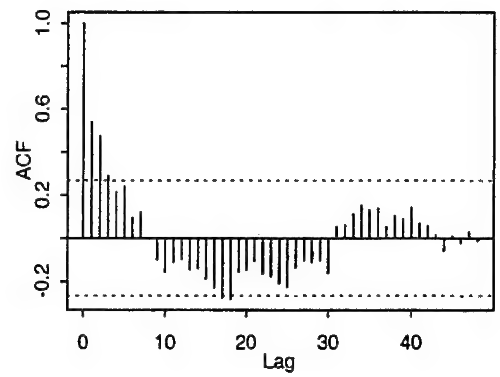
Magazine Ad Impressions



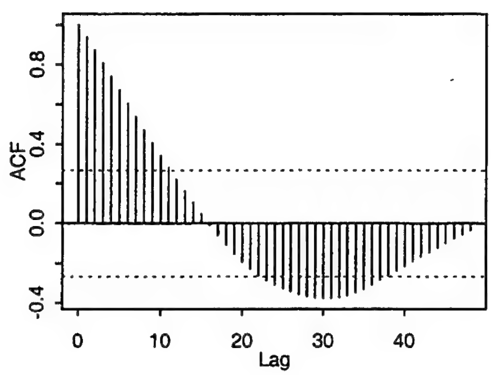
Newspaper Ad Impressions



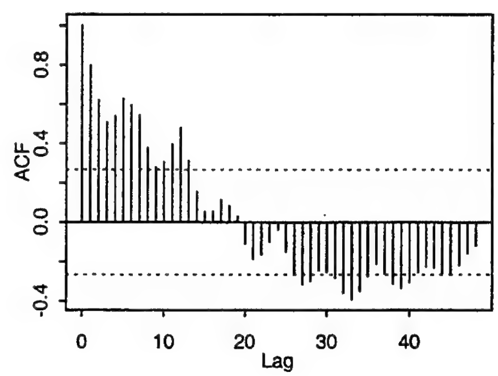
% Receiving College Fund



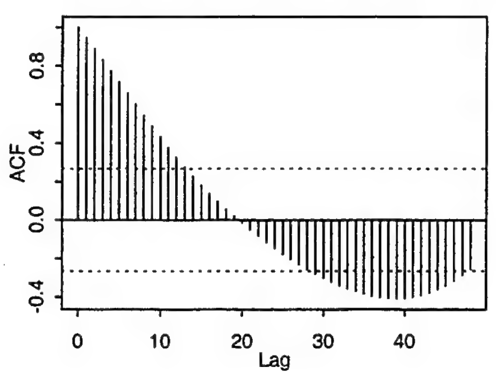
Target Population



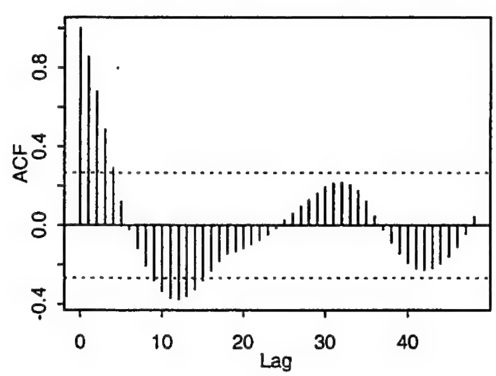
Unemployment Rate



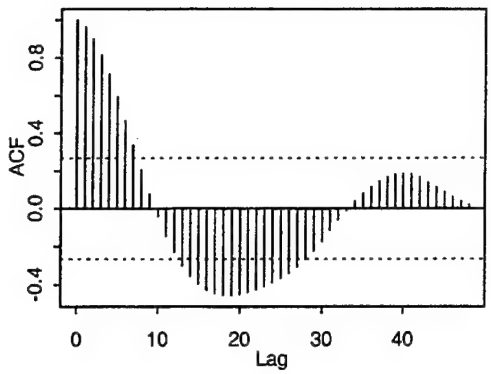
High School Grad Wage Level



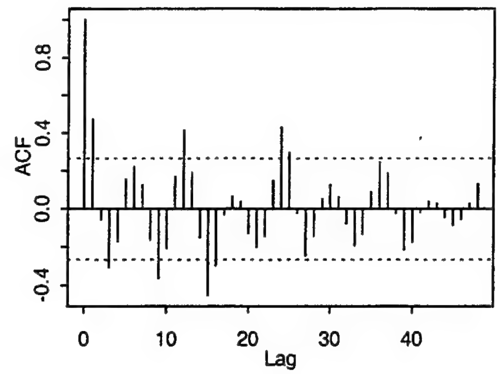
College Attendance Rate



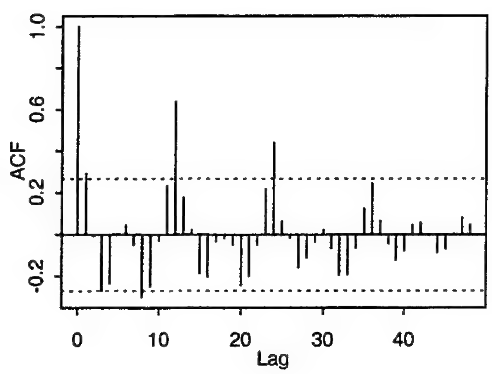
College Premium



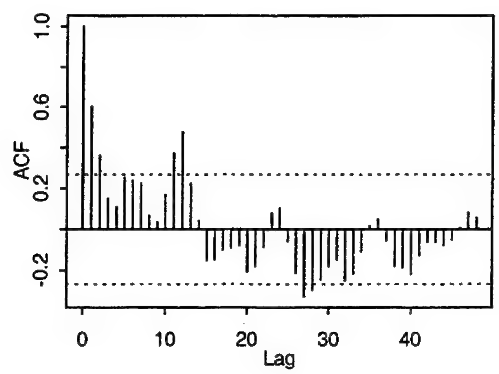
Air Force H-Q Male Rec Prod



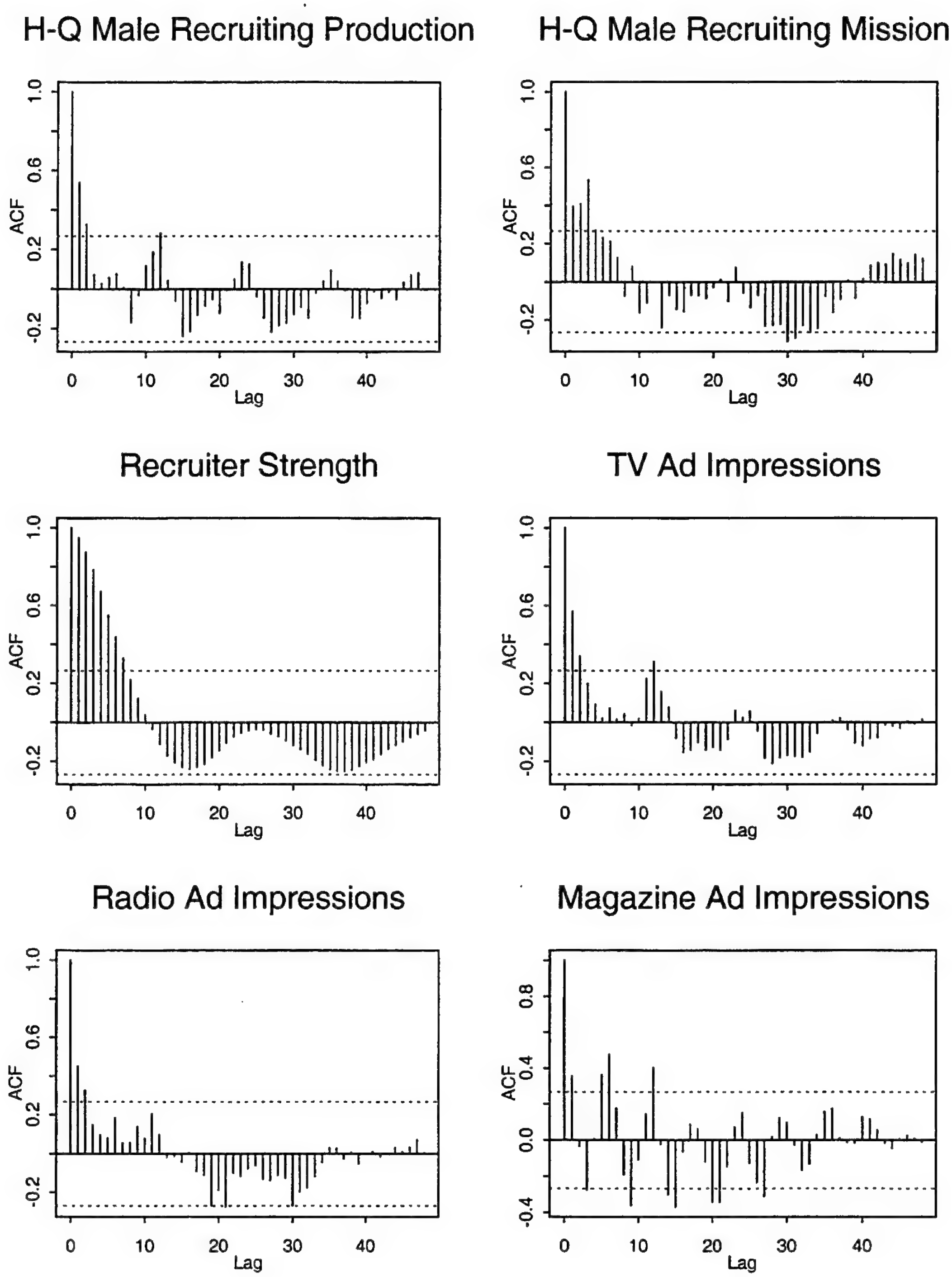
Marine Corps H-Q Male Rec Prod

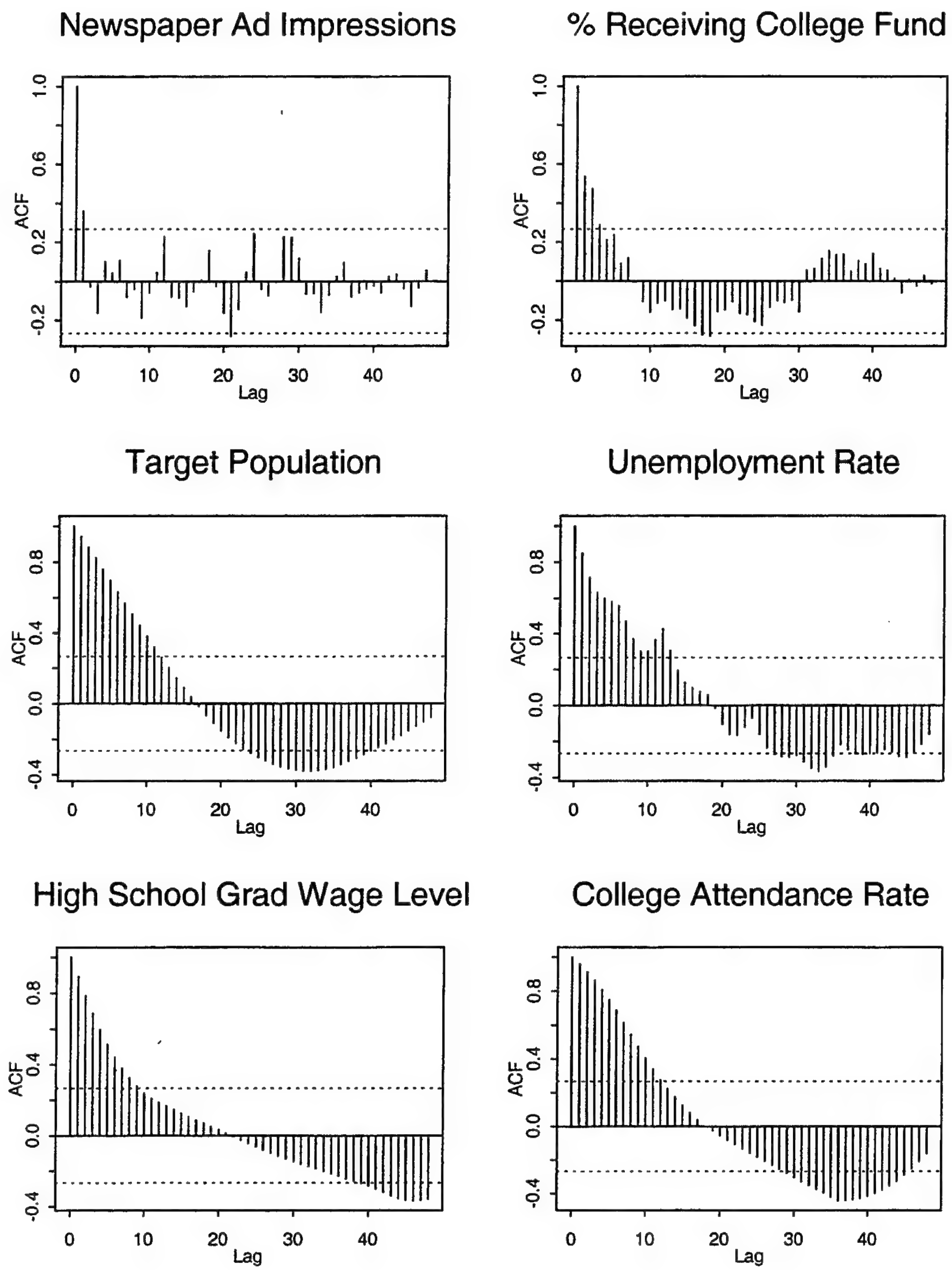


Navy H-Q Male Rec Prod

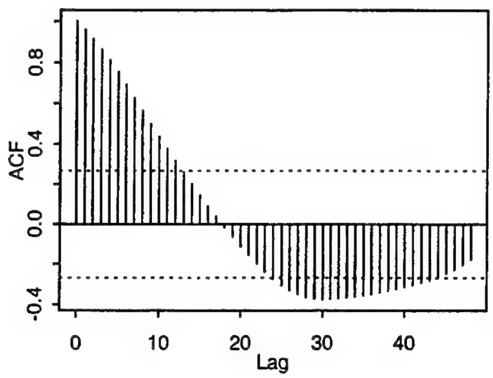


Fifth Brigade Autocorrelation Function Correlograms

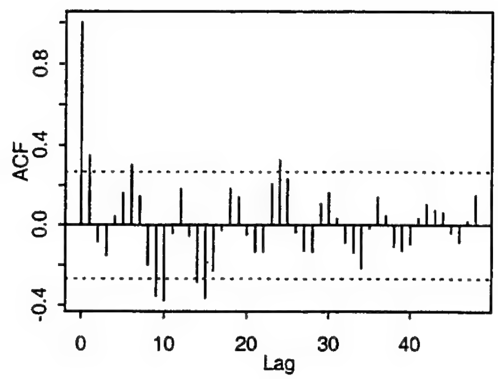



Newspaper Ad Impressions
% Receiving College Fund
Target Population
Unemployment Rate
High School Grad Wage Level
College Attendance Rate
ACF
Lag

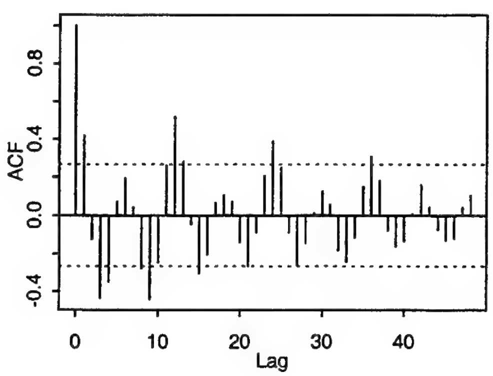
College Premium



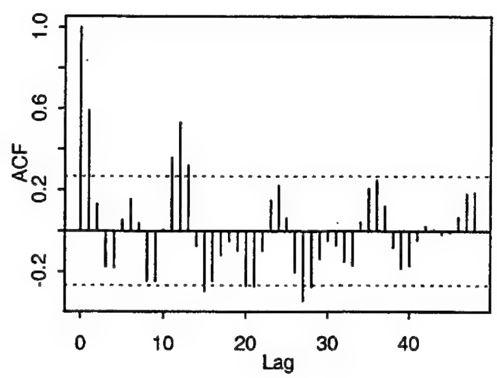
Air Force H-Q Male Rec Prod



Marine Corps H-Q Male Rec Prod



Navy H-Q Male Rec Prod



THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Asch, Beth J., and James Dertouzos, *Educational Benefits Versus Enlistment Bonuses*, Report MR-302-OSD, RAND, 1996.

Asch, Beth J., Kilburn, and Klerman, *Attracting College-Bound Youth Into the Military*, Report MR-984-OSD, RAND, 1999.

Bowerman, Bruce and O'Connell, *Forecasting and Time Series*, Duxbury Press, 1979.

Box, George and Gwilym Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, 1976.

Bureau of Labor Statistics, "Employment Status of the Noninstitutional Civilian Population 16 Years and Older." Available: ftp://ftp.bls.gov/pub/special.requests/lf/cpseea1.txt.

Bureau of Labor Statistics, "BLS Selective Data Access." Available: http://stat.bls.gov/sahome.html

Chatfield, Chris, *The Analysis of Time Series: An Introduction*, Chapman and Hall, 1996.

Cook, R.D. and Weisberg, S., *Applied Regression Including Computing and Graphics*, Wiley-Interscience, 1999.

Cotterman, Robert F., *Forecasting Enlistment Supply: A Time Series of Cross Sections Model*, Report R-3252-FMP, RAND, 1986.

Defense Manpower Data Center, "Recruiter Strength 1995-1999." Specialized Data Report produced by Richard Moreno, DMDC West.

Dertouzos, James N., and Thomas Chesnutt, *Recruiting Effects of Army Advertising*, Report R-3577-FMP, RAND, 1989.

Dickey, Connie, "Shinseki Devises Recruiting Strategy to Ensure Goals Are Met", ArmyLink News, 24 November, 1999. Available: http://www.ditc.mil/armylink/news

Efron, Bradley, and Robert Tibshirani, *An Introduction to the Bootstrap*, Chapman & Hall, 1998.

Eitelberg, Mark J., and Stephen Mehay, *Marching Toward the 21st Century: Military Manpower and Recruiting*, Greenwood Press, 1994.

Goldhaber, Dan D. *An Econometric Analysis of the Enlisted Goaling Model*, Center for Naval Analyses, 1999.

Harvey, Andrew C., *Time Series Models*, The MIT Press, 1994.

High, Gil, "Talking with the New Chief Of Staff, " Soldiers, August 1999, p 2-3.

Interview between C. Kaylor, Colonel, US Army, Chief of Military Strength Analysis and Forecasting Division, Office of the US Army Deputy Chief of Staff for Personnel, and the author,1 December 1999.

Lewis, James M., *An Examination of the Influence of Environmental Factors on Recruiting Category I-IIIA Males*, Master's Thesis in Operations Research, Naval Postgraduate School, Monterey, California, 1987.

Marquis, Susan, and Sheila Kirby, *Economic Factors in Reserve Attrition*, Report R-3686-1-RA, RAND, 1989.

Orvis, Bruce R., and Martin Gahart, *Quality-Based Analysis Capability for National Youth Surveys*, Report R-3675-FMP, RAND, 1989.

Orvis, Bruce R., Gahart, and Hosek, *Predicting Enlistment for Recruiting Market Segments*, Report N-2852-FMP, RAND, 1989.

Orvis, Bruce R., Sastry, and McDonald, *Military Recruiting Outlook: Recent Trends in Enlistment Propensity and Conversion of Potential Enlisted Supply*, Report MR-677-A/OSD, RAND, 1996.

Parlier, Gregory, *USAREC Program, Analysis, and Evaluation Directorate: An Overview,* presentation to operations research students 28 May, 1999 at the Naval Postgraduate School.

Telephone conversation between R. Pettit, Major, US Army, US Army Recruiting Command Program, Analysis, and Evaluation Directorate, and the author, week of 11 October 1999.

Shinseki, Eric, "Address to the Eisenhower Luncheon" speech to the 45th Annual Meeting of the Association of the United States Army, 12 October, 1999.

U.S. Army Recruiting Command Public Information Site. Available: www.usarec.army.mil.

Warner, John T. and B. Asch, "The Economics of Military Manpower", in *Handbook of Defense Economics*, edited by Keith Hartley and Todd Sandler, Amsterdam, 1995.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center 2
   8725 John J. Kingman Road, STE 0944
   Fort Belvoir, Virginia 22060-6218

2. Dudley Knox Library 2
   Naval Postgraduate School
   411 Dyer Road
   Monterey, California 93943-5101

3. Professor Samuel Buttrey 1
   Code OR/Sb
   Naval Postgraduate School
   Monterey, California 93943-5002

4. CPT Keith Hattes 1
   ODCSPER
   Attn: DAPE-MPA-RP
   300 Army Pentagon
   Washington, DC 20310-0300

5. Professor Thomas Lucas 1
   Code OR/Lt
   Naval Postgraduate School
   Monterey, California 93943-5002

6. Professor Lyn Whitaker 1
   Code OR/Wh
   Naval Postgraduate School
   Monterey, California 93943-5002

7. Major Mark Young... 1
   US Army Recruiting Command ATTN: PAE
   1307 Third Avenue
   Fort Knox, KY 40121

8. Captain Eric C. Burger 1
   5633A Folger Street
   Fort Knox, KY 40121-5000